

# Density-functional approaches to noncovalent interactions: A comparison of dispersion corrections (DFT-D), exchange-hole dipole moment (XDM) theory, and specialized functionals

Lori A. Burns, Álvaro Vázquez- Mayagoitia, Bobby G. Sumpter, and C. David Sherrill

Citation: *J. Chem. Phys.* **134**, 084107 (2011); doi: 10.1063/1.3545971

View online: <http://dx.doi.org/10.1063/1.3545971>

View Table of Contents: <http://jcp.aip.org/resource/1/JCPSA6/v134/i8>

Published by the American Institute of Physics.

## Additional information on J. Chem. Phys.

Journal Homepage: <http://jcp.aip.org/>

Journal Information: [http://jcp.aip.org/about/about\\_the\\_journal](http://jcp.aip.org/about/about_the_journal)

Top downloads: [http://jcp.aip.org/features/most\\_downloaded](http://jcp.aip.org/features/most_downloaded)

Information for Authors: <http://jcp.aip.org/authors>

## ADVERTISEMENT

**physicstoday**

**Comment on any  
*Physics Today* article.**

*Physics Today* / Volume 65 / July 2012  
Previous Article | Next Article  
**Measured energy in Japan**  
David von Seggern  
(vonneg@seismo.unr.edu) University of Nevada  
July 2012, page 10  
DIGITAL OBJECT IDENTIFIER  
<http://dx.doi.org/10.1063/PT.3.1619>  
The article by Thorne Lay and Hiroo Kanamori is an interesting one. It discusses the energy released by the 2011 earthquake in Japan. The authors estimate that the earthquake released about 100 megajoules of energy. This is a large amount of energy, but it is only a fraction of the energy released by a 100-megaton explosion. The authors also discuss the energy released by the 1964 Chilean earthquake. They estimate that this earthquake released about 100 megajoules of energy. This is also a large amount of energy, but it is only a fraction of the energy released by a 100-megaton explosion. The authors conclude that the energy released by earthquakes is much smaller than the energy released by nuclear weapons.

**Comment on this article**  
By the act of hitting a ball with a bat, one calculates the force energy to deliver the ball to its new location, but one must also take into account that the ball extended its energy release to that which became struck by the ball as its momentum ceased and passed energy to the struck item. Therefore the parameters of the damage extend into the future when the received energy to that pushed upon later becomes released in a new event. Perhaps calculations of one added that in while another's calculations did not. E.M.C.  
Written by Edgar McCarroll, 14 July 2012 19:59

# Density-functional approaches to noncovalent interactions: A comparison of dispersion corrections (DFT-D), exchange-hole dipole moment (XDM) theory, and specialized functionals

Lori A. Burns,<sup>1</sup> Álvaro Vázquez-Mayagoitia,<sup>2</sup> Bobby G. Sumpter,<sup>3</sup>  
and C. David Sherrill<sup>1,4,a)</sup>

<sup>1</sup>Center for Computational Molecular Science and Technology, School of Chemistry and Biochemistry,  
Georgia Institute of Technology, Atlanta, Georgia 30332-0400, USA

<sup>2</sup>Department of Chemistry, University of Tennessee, Knoxville, Tennessee 37996-1600, USA

<sup>3</sup>Computer Science and Mathematics Division and Center for Nanophase Materials Sciences, Oak Ridge  
National Laboratory, Oak Ridge, Tennessee 37831-6367, USA

<sup>4</sup>School of Computational Science and Engineering, Georgia Institute of Technology,  
Atlanta, Georgia 30332-0280, USA

(Received 22 November 2010; accepted 28 December 2010; published online 25 February 2011)

A systematic study of techniques for treating noncovalent interactions within the computationally efficient density functional theory (DFT) framework is presented through comparison to benchmark-quality evaluations of binding strength compiled for molecular complexes of diverse size and nature. In particular, the efficacy of functionals deliberately crafted to encompass long-range forces, *a posteriori* DFT+dispersion corrections (DFT-D2 and DFT-D3), and exchange-hole dipole moment (XDM) theory is assessed against a large collection (469 energy points) of reference interaction energies at the CCSD(T) level of theory extrapolated to the estimated complete basis set limit. The established S22 [revised in J. Chem. Phys. **132**, 144104 (2010)] and JSCH test sets of minimum-energy structures, as well as collections of dispersion-bound (NBC10) and hydrogen-bonded (HBC6) dissociation curves and a pairwise decomposition of a protein–ligand reaction site (HSG), comprise the chemical systems for this work. From evaluations of accuracy, consistency, and efficiency for PBE-D, BP86-D, B97-D, PBE0-D, B3LYP-D, B970-D, M05-2X, M06-2X,  $\omega$ B97X-D, B2PLYP-D, XYG3, and B3LYP-XDM methodologies, it is concluded that distinct, often contrasting, groups of these elicit the best performance within the accessible double- $\zeta$  or robust triple- $\zeta$  basis set regimes and among hydrogen-bonded or dispersion-dominated complexes. For overall results, M05-2X, B97-D3, and B970-D2 yield superior values in conjunction with aug-cc-pVDZ, for a mean absolute deviation of 0.41 – 0.49 kcal/mol, and B3LYP-D3, B97-D3,  $\omega$ B97X-D, and B2PLYP-D3 dominate with aug-cc-pVTZ, affording, together with XYG3/6-311+G(3df,2p), a mean absolute deviation of 0.33 – 0.38 kcal/mol. © 2011 American Institute of Physics. [doi:10.1063/1.3545971]

## I. INTRODUCTION

The pervasiveness of hydrogen bonding and London forces among biological<sup>1–3</sup> and supramolecular<sup>4–6</sup> entities has motivated persistent efforts<sup>7–36</sup> to develop reliable computational methodologies for noncovalent interactions (NCI). Extensive work has established the importance of high levels of electron correlation,<sup>37–39</sup> as embodied by coupled-cluster through perturbative triples, CCSD(T),<sup>40</sup> for the proper characterization of dispersion attractions, yet achieving this “gold-standard” of chemical accuracy<sup>41</sup> presents a formal cost of  $O(N^7)$ , where  $N$  is proportional to system size. Meanwhile, the attractively efficient density functional theory (DFT) [formally  $O(N^3) - O(N^4)$ , or  $O(N^5)$  for double-hybrids] can falter for even qualitative descriptions.<sup>42–45</sup> The introduction of the DFT+dispersion (DFT-D) (Refs. 16–19) and exchange-hole dipole moment (XDM) (Refs. 20–25) methods and the crafting of several improved functionals<sup>18,26–34</sup> has transformed the vista of

quantum chemical techniques available for nonbonded systems. These are among the promising approaches emerging from a very active field<sup>15–36</sup> that strives to improve performance for classes of NCI governed by disparate intermolecular forces and spanning broad length scales within the, as yet, inexact and nonsystematically refinable DFT environment. The present work has enlisted established and modern functionals to examine the current capabilities of density functional-based techniques, with special emphasis being directed toward experiential assessments of best practices.

Density-functionals developed for covalent systems are largely successful in treating hydrogen bonding and other electrostatically dominated noncovalent interactions that act over similarly short length scales ( $< 2$  Å).<sup>46</sup> In contrast, the dispersion attraction, which arises from correlated motions of electrons, is prominent chiefly over longer distances [ $\sim 2 - 5$  Å (medium-range) and  $> 5$  Å (long-range)], and none of the typical components of a functional, the local electron density ( $\rho_\sigma$ ), its gradient [ $\nabla\rho_\sigma$ , present in generalized gradient approximation (GGA) functionals], or its kinetic energy ( $\tau_\sigma$ , present in meta-GGA functionals), is fully capable of acting

<sup>a)</sup>Electronic mail: sherrill@gatech.edu.

over a suitable span. Hybrid functionals include long-range behavior (nonlocality) through Hartree–Fock exchange but remain local in correlation and, therefore, are also unable to correctly describe the  $R^{-6}$  asymptotic distance-dependence of dispersion forces. The consequent challenges to adapting affordable-scaling wavefunction and DFT methods for dispersion are amply reviewed in the recent literature;<sup>47–52</sup> Ref. 48 is particularly recommended for the nonspecialist and Ref. 52 for the specialist. Among the more empirical treatments developed are dispersion-corrected atom-centered potentials (DCACP) (Ref. 15) and DFT-D.<sup>16–19</sup> For common implementations of the former, a pseudopotential optimized to replicate reference interaction energies (IE) is placed over each atom, thereby incorporating medium-range correlation but not capturing the correct asymptotic form, while for the latter, a damped atom–atom dispersion term is appended to the total DFT energy, thereby accurately encoding the long-range attraction, yet necessitating a carefully tuned damping function to address double-counting of the correlation energy. XDM (Refs. 20–25) is constructed similarly to DFT-D as a correction to the base electronic energy but, by modeling the instantaneous dipole that arises between an electron and its exchange hole, generates dispersion coefficients that are aware of the chemical environment. Also active are efforts to include medium-range dispersion in conventional semilocal density functional theory<sup>18,26–30</sup> or to incorporate correlation components from wavefunction theory;<sup>31–34,53</sup> both approaches sometimes address long-range effects through an additional DFT-D-like term.<sup>18,30,32</sup> A more complex task is the development of explicitly nonlocal correlation functionals from first principles. Currently at the forefront are vdW-DF (Ref. 35) and VV09,<sup>36</sup> which do not yet regularly exceed the accuracy of more empirical methods.

Given the rapid generation of new dispersion-inclusive DFT techniques over the past five years, assessments have often addressed only a few methods at a time in validation of a new approach. As DFT+dispersion gained momentum in 2006, a study by Antony and Grimme<sup>54</sup> appraised the performance of the -D correction appended to GGA and hybrid functionals for nucleobase pairs and S22,<sup>55</sup> a diverse set of noncovalent complexes, and concluded that DFT-D achieves good accuracy of 0.5–1.0 kcal/mol for a range of NCI when employed with basis sets of polarized triple- $\zeta$  quality or larger and without correcting for basis set superposition error (BSSE). The following year, Jurečka and co-workers<sup>56</sup> examined six DFT-D approaches with S22 and reiterated the basis set requirement while noting that DFT-D is broadly superior to the underlying functional for nonbonded systems and often surpasses MP2/cc-pVTZ. The advent of several new dispersion-tuned functionals in 2008 led to findings by Zhao and Truhlar<sup>29</sup> that M05-2X and M06-2X are the best among 17 methods (none of which were DFT-D) for noncovalent interactions and by Chai and Head-Gordon<sup>30</sup> that  $\omega$ B97X-D is the best overall performer in comparison to DFT-D; both estimations are based upon S22 and thermochemical test sets. Two works by one of the authors<sup>37,57</sup> examined nucleobase pairs and dispersion-dominated potential energy curves (PEC) with DFT-D and M0N-2X and noted that the latter methods excel at stacking and near-range interactions but fail

to capture attractive forces at longer distances as DFT-D does. In the past year, a study by Riley and co-workers<sup>58</sup> compared TPSS-D and M06-2X for stacking, hydrogen-bonding,  $X-H \cdots \pi$ , and dispersion curves and reached similar conclusions, finding that the DFT-D approach is well-behaved for the first two interaction motifs and M06-2X is preferable for the middle two while sometimes distorting the shape of the potential curve. Several functionals emerged as “best” in the current year. In a work by most of the authors,<sup>59</sup> B97-D and XYG3 achieved record low errors for several PEC of weakly interacting systems, while similar examinations of hydrogen-bonding curves<sup>60</sup> found that  $\omega$ B97X-D was most reliable. After an extensive review, Riley and co-workers<sup>52</sup> reported that B97-D and  $\omega$ B97X-D are currently the most recommendable DFT-based methods for NCI, guidance with which this study ultimately agrees. A contrasting accord was reached between Grimme *et al.*,<sup>19</sup> who compared old (DFT-D2) and new (DFT-D3) variants of 11 functionals beyond the effects of BSSE for S22 and classes of fundamental biomolecules and concluded that B2PLYP-D is best overall, and Gráfová and co-workers,<sup>61</sup> who evaluated M06-2X and several DFT-D methods for curves of S22 complexes and concluded that although the accuracy of dispersion treatments still hinders the recommendation of any DFT technique as a “black-box” for biomolecules, the B2PLYP-D method is the most balanced approach. The latest appraisal of XDM by Kannemann and Becke<sup>62</sup> established that coupling with a GGA functional, PW86PBE-XDM, achieves excellent results (average error  $\sim 0.33$  kcal/mol) with far less empiricism than DFT-D and far fewer parameters than is typical of dispersion-modified functionals. Interestingly, Grimme *et al.*<sup>19</sup> collected the best results from the literature for the S22 set and posited a mean absolute deviation of 0.20–0.25 kcal/mol as a limit for general purpose methods. The present work aims to draw methods and test systems from many of these previous studies and through variety of computational techniques examined, quantity of systems considered, and range of molecular configurations sampled, offer a broad assessment of current capabilities.

One aspect stressed as early as 2004 (Ref. 63) and gaining recognition in the field is the importance of incorporating nonequilibrium molecular configurations into the development and assessment of methods for NCI, from both mathematical and physical considerations. For the former, it is apparent that a technique can, by generating a shifted or distorted dissociation curve, supply an accurate interaction energy for a reference minimum geometry while predicting very poorly the actual equilibrium position; instances of this may be found in most any large collection of potential energy curves.<sup>31,37,58</sup> Moreover, systems of biological and supramolecular interest support a complex organization with multifarious NCI, many of which are constrained into nonequilibrium geometries (e.g., stacking and interstrand contacts in DNA). Although hydrogen bonding is a highly directional attraction whose strength diminishes rapidly with improper alignment, the largely nondirectional nature of van der Waals (vdW) forces leads to their accumulation in sizable molecules. A database of very accurate potential energy curves focused on  $\pi$ -interactions ( $\pi$ - $\pi$ , S-H/ $\pi$ , and C-H/ $\pi$ )

TABLE I. Classification and performance of density functionals. In this and subsequent tables, functionals are ordered (within DFT, dispersion-including DFT, and XDM sections) by improving results for the overall S22 test set with the aug-cc-pVTZ basis. Methods exhibit, according to the symbols, satisfactory (✓) or excellent (\*) performance for purely hydrogen-bonded (HB), purely dispersion-bound (DD), and overall (TT) systems with double- $\zeta$  and triple- $\zeta$  basis sets. The MAD (kcal/mol) averaged over all test sets is also charted.

Density functional	Type	Perdew Ladder rung <sup>103</sup>	aug-cc-pVDZ				aug-cc-pVTZ			
			HB	DD	TT	MAD	HB	DD	TT	MAD
B97	GGA	2				4.08				4.50
BP86	GGA	2				3.04				3.50
B3LYP	hybrid GGA	4				2.81				3.28
B970	hybrid GGA	4				2.40				2.88
PBE	GGA	2				1.74				2.19
PBE0	hybrid GGA	4				1.67				2.11
B2PLYP	double-hybrid GGA	5	✓			0.98				1.53
XYG3	double-hybrid GGA	5	✓	✓	✓	0.58	✓ <sup>a</sup>	* <sup>a</sup>	* <sup>a</sup>	0.35 <sup>a</sup>
M05-2X	hybrid meta-GGA	4	✓	*	*	0.42				0.76
PBE0-D	hybrid GGA-D	4		*		0.73				0.68
PBE-D	GGA-D	2		✓		0.82				0.53
B970-D	hybrid GGA-D	4	*	*	*	0.41	*		✓	0.44
BP86-D	GGA-D	2				0.74		✓		0.52
B2PLYP-D	double-hybrid GGA-D	5				0.84	*	*	*	0.34
B97-D	GGA-D	2	*	*	*	0.49	*	*	*	0.38
B3LYP-D	hybrid GGA-D	4		✓		0.72	✓	*	*	0.37
M06-2X	hybrid meta-GGA	4	*		✓	0.50	✓	✓	✓	0.44
$\omega$ B97X-D	LC hybrid GGA-D	4				0.79	*	*	*	0.33
B3LYP-XDM	hybrid GGA-XDM	4	*	*	*	0.43	*	*	*	0.22

<sup>a</sup>Successful XYG3 results follow from the 6-311+G(3df,2p) basis set, not aug-cc-pVTZ.

has been collected by Sherrill *et al.*<sup>37</sup> Other efforts have been directed toward the S22 test set. Merz and co-workers<sup>64</sup> traced 14-point potential curves from MP2 and CCSD(T) calculations and derived analytical Morse potentials for most S22 systems to aid in force-field development. Recently, Hobza and co-workers<sup>61</sup> constructed limited, five-point PEC for the entire set at a higher level of theory and carefully analyzed the robustness of wavefunction and DFT approaches over the stages of dissociation. In a similar approach, Riley and co-workers<sup>58</sup> generated seven potential energy profiles spanning stacking, hydrogen-bonding, X-H $\cdots\pi$ , and dispersion interactions to examine functional performance with respect to range. A great many other high-quality dissociation curves have been offered in the literature. Systems at diverse geometries are better representative of target molecules and can provide a challenge for many computational methods, especially those that, supporting only midrange correlation, are expected to treat NCI only at near-equilibrium.

The demands of “interesting” systems for accurate and affordable computational treatments have prompted an assessment of DFT-based methods for noncovalent interactions. A coupled motivation is the aid of the nonspecialist for whom the proliferation of density functional schemes in the literature can make difficult the selection of the one best suited to his purposes. Following a description of the computational techniques employed (Sec. II A) and the chemical complexes examined (Sec. II B), DFT results will be presented and interpreted with regard to nonbonding

motifs (Secs. III B and III C), basis-set influence (Sec. III D), and computational efficiency (Sec. III E).

## II. COMPUTATIONAL METHODS

### A. DFT methods appraised

The techniques assessed in this work fall into three broad approaches for integrating dispersion into density functional theory: (i) functionals developed without special regard for NCI, yet which yield reasonable performance through *a posteriori* DFT-D correction; (ii) functionals constructed to include mid-/long-range interactions or optimized in the presence of a dispersion-correction term; and (iii) functionals utilized with the exchange-hole dipole moment method of obtaining dispersion coefficients. In this section, brief descriptions are given for these categories and the selected functionals, with a summary appearing in Table I. In the case of class (i), additional results will often be tabulated for the base functional to facilitate comparison.

#### 1. Semiempirical dispersion contribution

The local or semilocal character of conventional density functionals necessarily leads to neglect of the long-range correlation interactions which capture attractive van der Waals forces. Initially proposed by Wu and Yang<sup>16</sup> and assiduously developed by Grimme,<sup>17–19</sup> the DFT+Dispersion method appends to the base functional a scaled, damped, and fitted



leading term to the well-known dispersion energy series,  $E_{\text{disp}} = -C_6/R^6 - C_8/R^8 - C_{10}/R^{10} - \dots$ . Many calculations in this work follow the DFT-D2 (Ref. 18) variant implemented in Q-Chem, where the correction takes the explicit form

$$E_{\text{disp}}^{\text{D2}} = -s_6 \sum_{i,j>i}^{N_{\text{at}}} \frac{C_6^{ij}}{R_{ij}^6} f_{\text{damp}}(R_{ij}). \quad (1)$$

Here, dispersion coefficients,  $C_6^{ij}$ , obtained from the geometric mean of tabulated elemental values, are summed over interatomic distances,  $R_{ij}$ , modulated by a damping function,  $f_{\text{damp}}(R_{ij})$ , that gradually activates the dispersion correction over a distance characterized by the sum of the two atomic vdW radii, while an overall scaling term,  $s_6$ , is optimized to be unique to each  $E_{\text{xc}}$  functional. The scaling factors employed are  $s_6^{\text{B2PLYP}} = 0.55$ ,  $s_6^{\text{PBE0}} = 0.6$ ,  $s_6^{\text{PBE}} = 0.75$ ,  $s_6^{\text{B970}} = 0.75$ ,<sup>65</sup>  $s_6^{\text{BP86}} = 1.05$ ,  $s_6^{\text{B3LYP}} = 1.05$ , and  $s_6^{\text{B97}} = 1.25$ .

Grimme *et al.* recently presented a refined method, DFT-D3,<sup>19</sup> which incorporates an additional  $R^{-8}$  term in the dispersion series and adjusts the  $C_6^{ij}$  combination formula and damping function. The individual atomic  $C_6^i$  are interpolated from several reference values based upon coordination numbers extracted from the molecular structure, rather than assigned solely by atomic identity as in DFT-D2, and thereby incorporate some awareness of the chemical environment into an otherwise largely heuristic correction. The -D3 dispersion has the form

$$E_{\text{disp}}^{\text{D3}} = - \sum_{n=6,8} s_n \sum_{i,j>i}^{N_{\text{at}}} \frac{C_n^{ij}}{R_{ij}^n} f_{\text{damp}}(R_{ij}, s_{r,n}), \quad (2)$$

where  $s_{r,6}$  and  $s_8$  are the customary nonunity parameters fitted for individual functionals and tabulated in Ref. 66. In this work, both the DFT-D2 and DFT-D3 modifications have been applied to a number of density functionals described below.

**PBE-D and PBE0-D.** The Perdew–Burke–Ernzerhof (PBE) functional<sup>67</sup> is a widely used member of the generalized gradient approximation class, represented as  $E_{\text{xc}}^{\text{GGA}}[\rho] = E_{\text{xc}}^{\text{LDA}} + a_x \Delta E_x^{\text{GGA}} + a_c \Delta E_c^{\text{GGA}}$ , where the coefficients  $a_x$  and  $a_c$  that scale the GGA exchange and correlation corrections to the local density approximation (LDA) are customarily unity. Developed under the tenet that  $E_{\text{xc}}$  functionals should be anchored in physical principles, PBE was obtained from continuous extrapolation of the correlation-exchange energy of the homogeneous electron gas and is defined only in terms of fundamental constants (with the exception of its LDA component). PBE0 is a hybrid (*vide infra*) extension of PBE through addition of a measure ( $\chi_x = 0.25$ ) of HF exchange.<sup>68,69</sup>

**BP86-D.** An alternate approach for developing accurate  $E_{\text{xc}}$  functionals lies in establishing a mathematical form and fitting its parameters to reliable molecular training sets. In this manner, the Becke88 (Ref. 70) exchange functional ( $\Delta E_x^{\text{B}}$ ) encoded correct long-range asymptotic behavior, then fixed

its empirical constant from atomic data; its combination with the Perdew86 (Ref. 71) correlation correction ( $\Delta E_c^{\text{P86}}$ ) defines BP86.

**B3LYP-D.** An immensely popular functional well-validated for organic molecules, B3LYP (Refs. 72,73) follows the customary formula for hybrid GGAs,  $E_{\text{xc}}^{\text{hybrid GGA}}[\rho] = E_{\text{xc}}^{\text{LDA}} + \chi_x (E_x^{\text{exact}} - E_x^{\text{LDA}}) + a_x \Delta E_x^{\text{GGA}} + a_c \Delta E_c^{\text{GGA}}$ , where the last two terms represent the Becke88 (Ref. 70) ( $\Delta E_x^{\text{B}}$ ) and Lee–Yang–Parr<sup>74</sup> (LYP) ( $\Delta E_c^{\text{LYP}}$ ) gradient corrections to LDA for exchange and correlation, respectively, and the second term expresses the mixing of exact HF (Ref. 75) and approximate density-based exchange. The three semiempirical coefficients,  $\chi_x = 0.20$ ,  $a_x$ , and  $a_c$ , were optimized against selected experimental thermochemical data.

**B970-D.** Employing a purely empirical approach to  $E_{\text{xc}}$  construction, Becke proposed B97,<sup>76</sup> a hybrid ( $\chi_x = 0.1943$ ) power-series expansion of gradient correction terms with coefficients optimized to fit G2/97 (Refs. 77,78) thermochemical data. The name for the dispersion-corrected form of this functional, B97-D, has been supplanted by Grimme's recent reparameterization and dehybridization (*vide infra*).<sup>18</sup> As this latter version is expected to be more widely used in the current literature, the original functional will be denoted in this paper as B970 (B970-D for its DFT-D variant).

## 2. Specially designed functionals

**M05-2X and M06-2X.** Members of successive families of functionals developed by Truhlar and co-workers, M05-2X (Ref. 28) and M06-2X (Ref. 29) are highly parameterized hybrid meta-GGAs intended for main-group thermochemistry and noncovalent interactions. By including NCI systems in the fitting set and incorporating double the amount of nonlocal exchange (2X implies  $\chi_x = 0.56$  and  $0.54$ , respectively), the functionals strive to capture medium-range electron correlation, for reasonable results up to  $\sim 5$  Å separation of complexes.<sup>47,57,79</sup> At long range, the correlation functional falls off exponentially, lacking the nonlocality to support the characteristic  $R^{-6}$  dispersion decay that other functionals include natively<sup>35</sup> or artificially through a semiempirical -D term.

**B97-D.** In B97-D,<sup>18</sup> Grimme sought to develop an efficient (nonhybrid) GGA with a rational partition of control over distinct length-scales of interaction. The B97 functional form of Becke (Ref. 76) was reparameterized in the presence of the van der Waals term of Eq. (1) through fitting to thermochemical data and a set of noncovalently bound complexes. By selecting a large dispersion scaling factor ( $s_6 = 1.25$ ), the underlying density functional is restricted to shorter length scales while the damped -D term describes medium to longer range correlation effects, thereby avoiding double-counting at intermediate distances that can arise when a -D correction is appended to a standard density functional.

**$\omega$ B97X-D.** Chai and Head-Gordon have developed a family of long-range corrected (LC) functionals around Becke's B97 mathematical form,<sup>76</sup> of which  $\omega$ B97X-D (Ref. 30) is a single-hybrid member optimized in the presence of an empirical dispersion term proportional to Eq. (1). Imposing distinctive handling of exchange based on length-scales permits long-range interactions to be treated by HF exchange (essentially  $\chi_x = 1.00$ ) and short-range interactions to be treated by an effectively hybrid exchange functional ( $\chi_x = 0.22$ ). The presence of the -D term is found to shift upward the inter-electronic distances over which exact exchange is appropriate.

**B2PLYP-D.** Grimme formulated the first modern fifth-rung double-hybrid functional, B2PLYP (Ref. 31), according to the general form  $E_{xc}^{\text{double hybrid GGA}}[\rho] = E_{xc}^{\text{LDA}} + \chi_x(E_x^{\text{exact}} - E_x^{\text{LDA}}) + a_x \Delta E_x^{\text{GGA}} + \chi_c(E_c^{\text{PT2}} - E_c^{\text{LDA}}) + a_c \Delta E_c^{\text{GGA}}$ , wherein the GGA (specifically,  $\Delta E_x^{\text{B}}$  and  $\Delta E_c^{\text{LYP}}$ ) is corrected by both a large fraction of exact HF exchange ( $\chi_x = 0.53$ ) and a portion of MP2 double-excitation correlation ( $\chi_c = 0.27$ ). B2PLYP-D (Ref. 32) inherits the B2PLYP hybridization parameters which are fixed by optimization with respect to the G2/97 (Refs. 77 and 78) set and appends the van der Waals correction of Eq. (1) where the scaling parameter,  $s_6$ , has been refitted for best performance with the S22 test set. The relatively small value of  $s_6 = 0.55$  reflects the intended role of the -D term (in company with PT2 correlation) to address long-range dispersion effects, while the GGA functional covers the short-range regime of electron correlation, and PT2 subsumes medium-range system-sensitive components.

**XYG3.** Zhang *et al.*<sup>34</sup> recently developed a promising double-hybrid functional XYG3 based on the components of B3LYP,  $\Delta E_x^{\text{B}}$  and  $\Delta E_c^{\text{LYP}}$ . By permitting independent variation of the proportion of GGA exchange correction,  $a_x$ , as well as the two hybridization parameters, and optimizing them against the G3/99 (Ref. 80) set of thermochemical data, a functional with a large fraction of exact HF exchange ( $\chi_x = 0.8033$ ) and a sizable fraction of PT2 correlation ( $\chi_c = 0.3211$ ) is obtained. The Kohn–Sham orbitals utilized for the PT2 portion of the calculation emerge from the intermediate B3LYP density, which is expected to provide a far better foundation than the local  $E_{xc}$  potential upon which B2PLYP relies.

### 3. Exchange-hole dipole moment

A strategy to place dispersion corrections on a less-empirical footing than DFT-D is the “semiclassical” exchange-hole dipole moment model first formulated in 2005 by Becke and Johnson<sup>20</sup> (later placed on firmer theoretical foundations<sup>22–24</sup>) from the idea that the nonzero dipole moment of an exchange hole<sup>81</sup> can induce an instantaneous dipole moment, and thus dispersion interactions, between nonoverlapping systems. The method is notable for its first-principles calculation of dispersion coefficients  $C_n$  according to

$$C_{2n+4,ij}(\alpha_i, \alpha_j, \langle M_{1,\dots,n}^2 \rangle_i, \langle M_{1,\dots,n}^2 \rangle_j) \quad n = 1, 2, 3, \quad (3)$$

where  $\alpha_i$  is the effective atom-in-molecule polarizability for atom  $i$ , estimated from a Hirshfeld partitioning of the molecular density, and  $\langle M_\ell^2 \rangle_i$  is the atomic expectation value of the squared  $\ell$ -th multipole moment.<sup>21,23,82</sup> Combined with numerical prescriptions for a critical dimension  $R_c$  related to the van der Waals radius, as required by the damping function, XDM forms a wholly self-consistent and first-principles treatment of dispersion effects through

$$E_{\text{disp}}^{\text{XDM}} = - \sum_{n=6,8,10} \sum_{i,j>i}^{N_{\text{at}}} \frac{C_n^{ij}}{R_{ij}^n} f_{\text{damp}}(R_{ij}). \quad (4)$$

Although analogous to Eq. (1), the XDM dispersion formula incorporates more terms, apprehends more chemically aware  $C^{ij}$  parameters, and includes pairwise terms only over distinct segments (generally monomers) of a molecular complex. The correction term of Eq. (4) has been employed in conjunction with the relatively inexpensive B3LYP functional and various damping function refinements to be detailed in a forthcoming paper.<sup>83</sup>

## B. Benchmark sets

Several sets of binary complexes have been assembled that (i) strive to be typical of “real-world” nonbonding interactions, (ii) encompass a span of structural arrangements and intermonomer distances, and (iii) support high-level interaction energy benchmarks. Obtaining these reference values typically involves estimation of the complete basis set (CBS) limit either through independent extrapolation of the HF and CCSD(T) correlation energies,  $E_{\text{CCSD(T)}}^{\text{CBS}} = E_{\text{HF}}^{\text{CBS}} + E_{\text{corr CCSD(T)}}^{\text{CBS}}$ , or more affordably, through independent extrapolation of HF and MP2 correlation energies followed by application of a “coupled-cluster correction” at a small basis set,  $E_{\text{CCSD(T)}}^{\text{CBS}} \approx E_{\text{MP2}}^{\text{CBS}} + \Delta \text{CCSD(T)}$   $= E_{\text{HF}}^{\text{CBS}} + E_{\text{corr MP2}}^{\text{CBS}} + (E_{\text{corr CCSD(T)}}^{\text{small}} - E_{\text{corr MP2}}^{\text{small}})$ . In describing the benchmark values for the following test sets, specification of the many variants of CBS extrapolation will take the form CBS(HF; MP2; CCSD(T)), where each component method is substituted by the basis employed for the respective level of theory. Further modifiers include “H:”, whose arguments are the basis sets for a two-point Halkier *et al.*<sup>84</sup> extrapolation and “Δ:”, whose arguments are the “small basis” in a  $\Delta$ CCSD(T) correction. All CCSD(T)/CBS estimations utilized counterpoise (CP) -corrected interaction energies. Consult the supplementary materials for a tabulation of the exact CBS treatment for each test set member.<sup>85</sup> Summary IE quantities for each benchmark set are available in Table II. The following is not to be considered a comprehensive list of test sets available for NCI; others include the NCCE31 set<sup>27,86</sup> of Zhao and Truhlar and the BEGDB online database<sup>87</sup> maintained by Hobza's group.

**S22.** Jurečka *et al.* designed a compact test set representative of nonbonded interactions, denoted S22.<sup>55</sup> The set is constructed from balanced contributions by seven hydrogen-bonded (HB), eight dispersion bound (DD), and seven mixed

TABLE II. Test set statistics. Interaction energy means and, in parenthesis, ranges (kcal/mol) are presented for each test set and subcategory.

Test set	Hydrogen bonded		Mixed influence		Dispersion bound		Overall	
S22	-13.94	(-20.69, -3.15)	-3.87	(-7.09, -1.50)	-4.53	(-11.66, -0.53)	-7.31	(-20.69, -0.53)
NBC10/HBC6	-10.76	(-26.29, -0.12)	-1.38	(-2.95, +9.34)	-1.17	(-2.87, +3.52)	-4.96 <sup>b</sup>	(-26.29, +9.34)
NBC10/HBC6 <sup>a</sup>	-18.18	(-26.29, -15.71)	-2.73	(-2.95, -2.50)	-1.92	(-2.87, -0.42)	-7.91 <sup>b</sup>	(-26.29, -0.42)
HSG	-13.11	(-18.98, -7.53)	-3.94	(-6.28, -0.55)	-1.00	(-2.18, +0.39)	-4.14	(-18.98, +0.39)
JSCH	-20.39	(-51.40, -5.20)	-1.13	(-5.20, +3.10)	-7.84	(-20.35, +2.45)	-9.95	(-51.40, +3.10)

<sup>a</sup>For each dissociation curve, includes only five points contiguous to minimum.<sup>b</sup>Average performed over both HBC and NBC test sets.

influence (MX) complexes varying in size from very small (e.g., water dimer) to substantial (e.g., adenine · thymine complex). As S22 was intended for training approximate methods, the authors took care to evaluate these systems at a robust level of theory; however, advances have enabled Takatani *et al.*<sup>88</sup> to reassess the published interaction energies, thereby correcting some values by up to 0.6 kcal/mol. The present work retains the molecular geometries from the original S22 formulation while using the revised reference energies of Takatani (formally designated S22A). The CBS estimation procedure was performed for the smaller systems through a direct extrapolation of the CCSD(T) correlation energy, CBS(aQ; - ; H:aT-aQ), and for the larger systems with the conventional two-step approach, CBS(aQ; H:aT-aQ; Δ:H:aD-aT). See Sec. II C for details of basis set abbreviations. Recent work by Gráfová *et al.*<sup>61</sup> extended the scope of the S22 set by constructing partial dissociation curves (S22 × 5) for each of its members, meanwhile adjusting the interaction-type categorization of complexes on the basis of symmetry-adapted perturbation theory (SAPT), yet leaving unchanged the original benchmark level of theory. Although the authors agree with the reapportionment, particularly the exchange of entries 13 and 20, the present work retains the original divisions for consistency with other test sets and literature values. S22 comprises 22 molecular systems at independent (near-equilibrium) geometries.

**NBC10.** Many previously published potential energy curves of nonbonded systems with frozen-geometry monomers have been collected into a test set named *Nonbonded Curves* or *NBC10*. Members include the sandwich (Bz·Bz-S), T-shaped (Bz·Bz-T), and 3.2, 3.4, and 3.6 Å-separated parallel-displaced (Bz·Bz-PD32, Bz·Bz-PD34, and Bz·Bz-PD36) configurations of the benzene dimer,<sup>37</sup> benzene-hydrogen sulfide (Bz·H<sub>2</sub>S),<sup>37</sup> benzene-methane (Bz·Me),<sup>37</sup> methane dimer (Me·Me),<sup>89</sup> the antiparallel sandwich (Py·Py-S2) pyridine dimer,<sup>90</sup> and a T-shaped (Py·Py-T3) pyridine dimer.<sup>90</sup> Benchmark interaction energies for these species typically have been computed at the CBS(aQ; H:aT-aQ; Δ:H:haD-haT) level of theory (sometimes Δ:H:aD-aT), with the exception of the methane dimer curve at CBS(aQ; - ; H:aT-aQ), which results from a direct extrapolation of CCSD(T) correlation energies, and the T-shaped pyridine dimer at CBS(aQ; H:aT-aQ; Δ:aD), which utilizes a lower level ΔCCSD(T) correction. The last is not expected to undermine the quality of the benchmark

values since T-shaped complexes have been found to be less sensitive than other configurations to the size of the small basis set employed in the coupled-cluster correction. When evaluating the performance of NBC10 over the broad categories established for the S22 test set, the ten curves are partitioned into Dispersion Bound ⊇ (Bz·Bz-S, Bz·Bz-PD, Bz·Me, Me·Me, Py·Py-S2) and Mixed Influence ⊇ (Bz·Bz-T, Bz·H<sub>2</sub>S, Py·Py-T3) subsets. NBC10 comprises ten molecular systems at 184 geometries.

**HBC6.** Recent efforts to expand the high-quality reference material available for hydrogen bonding systems across varied geometries has led to the collection of a test set denoted *Hydrogen-Bonded Curves* or *HBC6*.<sup>60</sup> The set is composed of formic acid (FaOO), formamide (FaON), and formamidine (FaNN) monomers in both homogeneous (FaOO·FaOO, FaON·FaON, FaNN·FaNN) and mixed (FaOO·FaON, FaON·FaNN, FaOO·FaNN) complexes. Together, these encompass many of the most common hydrogen bonding patterns, including O··H··O, O··H··N, and N··H··N. The benchmark interaction energies along these curves have been computed uniformly at a counterpoise-corrected CBS(aQ; H:aT-aQ; Δ:H:haD-haT) level of theory. In contrast to NBC10, HBC6 curves are fully optimized; that is, relaxed monomers are employed. Available potential energy profiles which additionally correct for deformation energy are not considered here. HBC6 comprises six molecular systems at 118 geometries.

**HSG.** Merz and co-workers examined an HIV-II protease protein with a bound indinavir ligand and decomposed the docking site into pairs of chemical fragments designated *HSG*.<sup>91</sup> The resulting complexes, which are not necessarily equilibrium configurations, contain between 19 and 32 atoms and are partitioned into the customary HB, MX, and DD subsets, as guided by the polarity of the amino acid side chain and decompositions from symmetry-adapted perturbation theory. Reference interaction energies are available at the CBS(aQ; H:aT-aQ; Δ:H:haD-haT) level of theory. HSG comprises 21 molecular systems at independent geometries.

**JSCH.** Jurečka *et al.* also constructed a large test set of 124 nucleobase pairs denoted *JSCH*.<sup>55</sup> In parallel to the HB, MX, and DD subsets of S22, JSCH includes three broad geometrical arrangements: 38 hydrogen-bonded (coplanar bases),



32 interstrand (adjacent base pairs on different strands), and 54 stacked (adjacent base pairs on the same strand) complexes. Efforts to evaluate the interaction energies for this set at a tractable level of theory led to neglect of the  $\Delta\text{CCSD(T)}$  correction for some hydrogen-bonded systems and the use of less-intensive basis sets for an estimated error of up to 0.8 kcal/mol. The published interaction energies are generally of CBS(H:aD-aT; H:aD-aT;  $\Delta$ :6-31G\*(0.25)) quality, though select members are evaluated at H:aT-aQ for the MP2 segment. Note that reordering of several systems was required to bring the published geometries in line with the original authors' intentions.<sup>92</sup> JSCH comprises 124 molecular systems at independent geometries.

### C. Technical details

DFT methods were evaluated primarily with the double- $\zeta$  (D $\zeta$ ) and triple- $\zeta$  (T $\zeta$ ) basis sets of Dunning augmented by diffuse functions on all atoms, namely, aug-cc-pVDZ (aDZ) and aug-cc-pVTZ (aTZ).<sup>93,94</sup> Hybrid sets where cc-pVXZ (XZ) are placed on hydrogen and the corresponding aXZ on nonhydrogen atoms are denoted heavy-aug-cc-pVDZ (haDZ) and heavy-aug-cc-pVTZ (haTZ). Additional calculations for XYG3 were performed with 6-311+G(3df,2p) (pp2) since this basis set was used during the functional's development and is expected to provide the best performance. Note that in basis set specification for CBS extrapolations, the uninformative "Z" is dropped to conserve space (e.g., aD $\leftarrow$ aDZ). DFT results are not counterpoise corrected for basis set superposition error by the method of Boys and Bernardi<sup>95</sup> unless explicitly indicated.

The B97-D, XYG3, and B2PLYP-D methods have been accessed through the quantum chemistry code NWChem 5.1.1,<sup>96</sup> the former two locally implemented as described previously.<sup>59</sup> The remaining functionals, including all DFT-D2 and XDM calculations, were performed within Q-Chem 3.2.<sup>97</sup> Dispersion corrections to form DFT-D3 methods were obtained from the computer program made available by Grimme.<sup>66</sup> Ongoing refinement of the B3LYP-XDM technique has employed the S22 and HSG test sets, as well as major subsets of NBC10 and HBC6, as training sets for parameter optimization; thus, only JSCH results are completely unbiased. Timing results from the Q-Chem program were generated on a Dell 380 system featuring Pentium-D processors (3.0 GHz, single socket dual core, 1024 KB L2) and 4 GB memory. NWChem was executed on Cray XT5 hardware (National Institute for Computational Sciences Kraken and Jaguar) using a single node of AMD Opteron 2435 processors (2.6 GHz, Istanbul dual hexa-core) and 16 GB DDR2-800 memory. The system used for comparing timings is S22-7, a Watson-Crick-bonded adenine-thymine complex of  $C_1$  symmetry.

To ensure reliable interaction energies and, particularly, to obtain smooth curves for HBC6 and NBC10 sets with meta-GGA functionals,<sup>98</sup> a fine integration grid was used, employing 100 radial shells with 302 angular points in Q-Chem and 100 radial shells with 1202 angular points in NWChem. Truncation of atomic orbitals due to linear dependency was

employed according to default settings. Although this often leads to unmatched orbital counts between the dimer and monomers, the resulting interaction energy is estimated to deviate by significantly less than 0.01 kcal/mol for equilibrium structures. For dissociation curves, where discontinuities are both readily detected and more likely to occur (especially between frozen monomers and compact dimer geometries), adjustments to the dependency tolerance were taken as needed.

The performance of a method and basis set (together, a theoretical model chemistry<sup>99</sup>) for a given test set is evaluated by comparing the interaction energy (always in kcal/mol) for each component system to that of its reference CCSD(T)/CBS value and is summarized by the quantity mean absolute deviation ( $\text{MAD} = \frac{1}{n} \sum_{\text{sys}} |\text{DFT}_{\text{sys}} - \text{REF}_{\text{sys}}|$ ). A common measure of relative error, the mean absolute percent deviation ( $\text{MA\%D} = 100 \cdot \frac{1}{n} \sum_{\text{sys}} \left| \frac{\text{DFT}_{\text{sys}} - \text{REF}_{\text{sys}}}{\text{REF}_{\text{sys}}} \right|$ ), offers the advantage of being an appropriate comparison between complexes with binding energies of different magnitude, for instance those in HB and DD classes (cf. Table II). However, MA%D is unsuited for test sets involving potential energy curves along a dissociation coordinate, as very small absolute errors are present at large (and specific near) intermonomer distances, yielding exaggerated percentage errors. This may be seen in Fig. 1 where error profiles are plotted for a trial curve displaced vertically by 0.01 from a reference Lennard-Jones potential of depth 1 kcal/mol. Whereas the MAD (computed over a generous range of 0.8 – 2.4 the equilibrium dissociation) is constant, as per construction, the MA%D blows up wherever the reference approaches zero, effectively tracking the topology of the reference far more clearly than it reflects the error of interest, and amounts to 21% from a very minor

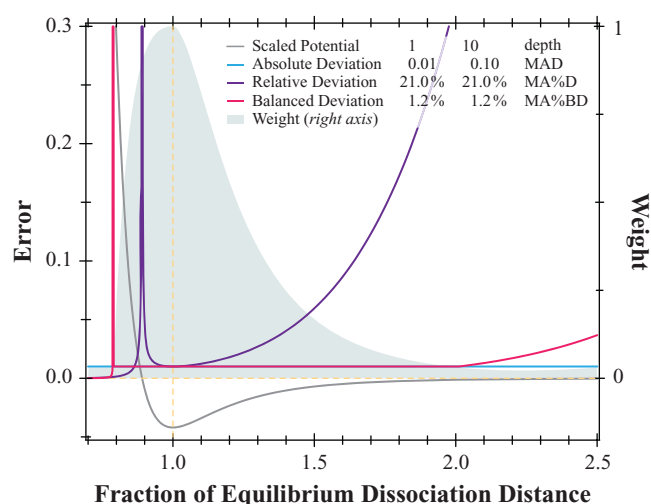


FIG. 1. Measures of error for a model dissociation curve. A Lennard-Jones potential (shown scaled) and its duplicate, displaced vertically by 1% of its well depth, are compared according to absolute, relative, and balanced definitions of deviation. Their summary statistics MAD, MA%D, and MA%BD, respectively, are computed over the chemically relevant portion of the curve [0.8 – 2.4 ·  $R_{eq}$ ] for potentials of depth one and ten. The graph demonstrates the difficulties in employing MA%D for nonequilibrium geometries and a possible solution in the form of balanced error, which follows the profile of the absolute error while being readily comparable over ranges of interaction energy.



displacement. The relative deviation for a curve of discrete points is strongly influenced by the chance that a reference geometry has an interaction energy near zero (note that the discontinuity in the head broadens with the magnitude of error). An alternate measure of relative error, denoted as balanced deviation or MA%BD for the summary statistic, is proposed that combines a number of desirable properties.

The balanced deviation is composed of an importance function and an error function. Through the former, points are weighted in proportion to their (reference) fraction of the equilibrium interaction energy,  $\text{REF}_{\text{eq}}$ , diminishing toward a nonzero minimum contribution,  $m$ , as the tail of the potential approaches zero and as the head approaches a multiple,  $p$ , of  $\text{REF}_{\text{eq}}$ . The error function over the head of the curve effectively shifts the zero-line discontinuity beyond the range of interest; for the tail, it reduces to the conventional definition of relative error. The product of these becomes  $\text{BD} = \frac{\text{DFT}_{\text{sys}} - \text{REF}_{\text{sys}}}{|\text{REF}_{\text{eq}}|}$  over essentially all important regions of a dissociation potential and is equivalent to MA%D for equilibria. Consult the supplementary material for a full definition.<sup>85</sup> Referring back to Fig. 1, the balanced deviation trace is closely attuned to the profile of the absolute error and is resistant to the mathematical irritants of the relative error. Comparing summary statistics for model potentials with depths differing by an order of magnitude, the MA%BD retains the relative scaling of MA%D while yielding a value of 1.2%, which far more accurately reflects the  $1\% \cdot \text{REF}_{\text{eq}}$  by which both are displaced. Relying on only two reasonably insensitive parameters and requiring only  $\text{REF}_{\text{eq}}$  and placement in the curve head or tail as additional knowledge for computation, the balanced deviation is a practical error quantity that warrants further consideration as dissociation curves are employed more heavily for parameterizing and evaluating NCI methods.

The absolute measure of error, MAD, is used primarily in this work. Nevertheless, MA%D counterparts to Tables III–VII and a MA%BD counterpart to Table IV (MA%D and MA%BD differ only for curves) are furnished in the supplementary material.<sup>85</sup> Other metrics, including minimal and maximal deviation ( $\text{minD} = \min_{\text{sys}} |x_{\text{sys}}|$  and  $\text{maxD} = \max_{\text{sys}} |x_{\text{sys}}|$ ), signed and unsigned mean deviations ( $\text{MD} = \frac{1}{n} \sum_{\text{sys}} x_{\text{sys}}$  and  $\text{MAD} = \frac{1}{n} \sum_{\text{sys}} |x_{\text{sys}}|$ ), and root-mean-square deviation ( $\text{RMSD} = \sqrt{\frac{1}{n} \sum_{\text{sys}} (x_{\text{sys}})^2}$ ), where  $x_{\text{sys}}$  represents one of absolute ( $x_{\text{sys}} = \text{DFT}_{\text{sys}} - \text{REF}_{\text{sys}}$ ), relative ( $x_{\text{sys}} = \frac{\text{DFT}_{\text{sys}} - \text{REF}_{\text{sys}}}{|\text{REF}_{\text{sys}}|}$ ), or balanced ( $x_{\text{sys}} \approx \frac{\text{DFT}_{\text{sys}} - \text{REF}_{\text{sys}}}{|\text{REF}_{\text{eq}}|}$ ) error quantities, are also available in supplementary material. Where applicable, statistics for subgroups (i.e., hydrogen bonded, mixed influence, dispersion bound) of test sets are collected. Hobza and co-workers have developed a series of metrics,  $\text{RQ}_1 - \text{RQ}_5$ , to rigorously evaluate the accuracy of a computational method for test sets structured into classes of interaction and dissociation curves.<sup>61</sup> For comparative purposes, the most universally applicable of these,  $\text{RQ}_1$ , which establishes strict (10%) and loose (20%) criteria for the overall relative error, is referenced in this work, substituting MA%BD for MA%D. It is readily gleaned that underbinding errors are positive and overbinding are negative.

## D. Reading strip charts

Results in this work can be conveniently illustrated by strip charts as in Figs. 2–4, the conventions for which are outlined here. Each horizontal strip represents the results for a single test set with a given model chemistry. Thin vertical lines plot the error in interaction energy (kcal/mol) for each member of the test set in either the underbound (right of the zero-error line) or overbound (left of the zero-error line) sector of the chart. Individual quantities lying beyond the range of the graph are omitted without annotation. The vertical lines are colored to indicate that the member belongs to the hydrogen-bonded (red), mixed-influence (green), or dispersion-bound (blue) subsets. The same coloring convention applies to markers indicating the MAD for the overall test set (black rectangles) and each of its subsets (colored squares), with their position being fixed in the “overbound” sector of the graph for the ease of comparison with the zero line, regardless of the preponderance of individual subset markings. For test sets composed of potential energy curves (i.e., NBC10 and HBC6), only five points centered on the equilibrium geometry of each curve are plotted and used to compute the MAD marker position; this strategy avoids the visual confusion that would arise from the many small-error points at the tail of a curve that are not directly comparable to the minimum-energy values in other test sets.

## III. RESULTS AND DISCUSSION

### A. Aspects of Grimme's -D correction

Conventional DFT methods applied to noncovalently bound systems typically return interaction energies that are extremely underbound, often possessing errors far in excess of +1 kcal/mol. The B97 functional underlying B97-D is particularly poor since it was designed to be used with its dispersion component. The DFT-D technique, which necessarily deepens the interaction well, is generally regarded as an advantageous or at least a benign correction to the DFT energy. While this guideline has few exceptions among dispersion-dominated and mixed-influence systems (B2PLYP-D/aDZ for HSG being a notable example), more ambiguous results are found for hydrogen-bonded complexes. With the aug-cc-pVDZ basis set, the PBE, PBE0, and B2PLYP functionals nearly consistently yield a lower MAD for the HB subset than their DFT-D2/DFT-D3 variants. This arises because underbound DFT values are overcorrected by DFT-D to form overbinding errors of greater magnitude than their uncorrected counterparts. The same trend is present in aug-cc-pVTZ, though to a lesser degree, so that the underbound DFT errors are balanced or slightly greater in magnitude than the overbound DFT-D deviations. The PBE-D and PBE0-D functionals are not recommended with either basis set for hydrogen-bonded systems, while B2PLYP-D/aug-cc-pVTZ may be used, as its -D term appears strategically scaled to achieve good results.

Grimme *et al.*'s refined DFT-D3 method<sup>19</sup> is far more attuned to the chemical environment than the widely used

TABLE III. Mean absolute deviation of the interaction energy (kcal/mol) for the S22 test set with noncounterpoise-corrected DFT methods.

Method and basis set	Hydrogen bonded		Mixed influence		Dispersion bound		S22	
	-D2	-D3	-D2	-D3	-D2	-D3	-D2	-D3
<b>aug-cc-pVDZ</b>								
B97	3.79		3.32		6.75		4.71	
BP86	1.81		2.70		5.44		3.42	
B3LYP	1.52		2.39		5.31		3.17	
B970	1.77		1.80		4.39		2.73	
PBE	0.63		1.42		3.62		1.97	
PBE0	0.51		1.13		3.33		1.74	
B2PLYP	0.64		0.64		1.79		1.06	
XYG3	0.49		0.29		0.57		0.46	
M05-2X <sup>a</sup>	0.45	0.33	0.43	0.95	0.21	1.06	0.36	0.79
PBE0-D	1.27	1.34	0.74	0.94	0.25	0.69	0.73	0.98
PBE-D	1.41	1.04	0.91	0.62	0.87	0.45	1.05	0.69
B970-D	0.25		0.57		0.24		0.35	
BP86-D	1.03	1.20	0.56	0.82	0.85	1.62	0.82	1.23
B2PLYP-D	0.84	0.71	1.09	0.98	1.50	1.30	1.16	1.01
B97-D	0.54	0.30	0.66	0.63	0.73	0.52	0.65	0.48
B3LYP-D	1.31	1.06	0.88	0.67	0.97	0.66	1.05	0.79
M06-2X <sup>a</sup>	0.38	0.24	0.50	0.79	1.14	1.57	0.69	0.90
$\omega$ B97X-D	0.64		0.74		1.23		0.89	
B3LYP-XDM	0.42		0.27		0.50		0.40	
<b>6-311+G(3df,2p)</b>								
XYG3	0.67		0.27		0.27		0.39	
B3LYP-D	0.92	0.66	0.70	0.49	0.66	0.34	0.75	0.49
<b>aug-cc-pVTZ</b>								
B97	4.13		3.75		7.39		5.19	
BP86	2.20		3.08		6.13		3.91	
B3LYP	1.95		2.79		5.91		3.66	
B970	2.15		2.29		5.11		3.27	
PBE	1.05		1.86		4.34		2.50	
PBE0	0.84		1.65		4.14		2.30	
B2PLYP	0.93		1.24		2.75		1.69	
XYG3	0.49		0.45		1.66		0.91	
M05-2X <sup>a</sup>	0.61	0.45	0.40	0.49	0.89	0.25	0.64	0.39
PBE0-D	0.89	0.96	0.31	0.42	0.62	0.31	0.60	0.55
PBE-D	0.98	0.61	0.46	0.22	0.19	0.47	0.53	0.44
B970-D	0.23		0.28		0.72		0.43	
BP86-D	0.67	0.88	0.24	0.44	0.32	0.97	0.41	0.77
B2PLYP-D	0.55	0.42	0.46	0.35	0.55	0.36	0.52	0.38
B97-D	0.76	0.48	0.37	0.34	0.11	0.27	0.40	0.36
B3LYP-D	0.88	0.63	0.47	0.26	0.41	0.22	0.58	0.36
M06-2X <sup>a</sup>	0.54	0.36	0.27	0.33	0.22	0.63	0.34	0.45
$\omega$ B97X-D	0.27		0.30		0.42		0.33	
B3LYP-XDM	0.17		0.26		0.15		0.19	

<sup>a</sup>Column labeled -D2 (-D3) contains figure for M0N-2X (M0N-2X-D3) functional variant.

DFT-D2, yet also disrupts the attractive simplicity of the earlier scheme. In comparing the two levels, -D3 is clearly inferior to its predecessor for the BP86 functional, with an increase in error ( $\text{MAD}^{\text{DFT-D3}} - \text{MAD}^{\text{DFT-D2}}$ ) of +0.24 kcal/mol averaged over all basis sets and test sets. The PBE and PBE0 methods, for which -D terms are of limited use (*vide supra*), show inconsistent results, as PBE-D3 and PBE0-D2 are preferred with aug-cc-pVDZ and neither variation is clearly favored with aug-cc-pVTZ. In contrast, B3LYP, B2PLYP, and B97 are unreservedly improved with the new -D3 correction by average MAD changes for HB, MX, DD and overall subsets of  $-0.17$ ,  $-0.02$ ,  $-0.11$ , and

$-0.11$  kcal/mol, respectively. The mixed-influence category irregularly supports -D2 or -D3, hence the small (signed) average improvement.

In very recent (unpublished) work,<sup>66</sup> Grimme offers a DFT-D3 variant for Truhlar's M05 and M06 families. Since M05-2X and M06-2X already incorporate some medium-range interactions, only the (long-range)  $R^{-6}$  term of Eq. (2) is retained. Compared to the underlying functionals, application of the recommended formula is detrimental to error patterns with the aug-cc-pVDZ basis set, yet beneficial with aug-cc-pVTZ. These findings are difficult to interpret since the circumstances of the parameters' optimization

TABLE IV. Mean absolute deviation of the interaction energy (kcal/mol) for the NBC10 and HBC6 test sets with noncounterpoise-corrected DFT methods.

Method and basis set	HBC6		NBC10					
	Hydrogen bonded		Mixed influence		Dispersion bound		NBC	
	-D2	-D3	-D2	-D3	-D2	-D3	-D2	-D3
<b>aug-cc-pVDZ</b>								
B97		2.33		2.51		3.82		3.42
BP86		0.98		2.12		3.36		2.98
B3LYP		0.82		2.05		3.40		2.99
B970		0.91		1.52		2.58		2.26
PBE		0.38		1.18		2.07		1.80
PBE0		0.42		1.05		2.02		1.72
B2PLYP		0.38		0.66		1.16		1.00
XYG3		0.55		0.44		0.45		0.45
M05-2X <sup>a</sup>	0.40	0.62	0.13	0.60	0.25	0.72	0.21	0.68
PBE0-D	1.29	1.37	0.33	0.62	0.22	0.63	0.25	0.63
PBE-D	1.25	1.05	0.54	0.48	0.61	0.60	0.59	0.56
B970-D	0.41		0.22		0.16		0.18	
BP86-D	0.94	1.06	0.50	0.66	0.54	1.02	0.52	0.91
B2PLYP-D	0.55	0.51	0.60	0.59	0.81	0.77	0.75	0.72
B97-D	0.51	0.39	0.35	0.45	0.65	0.53	0.56	0.51
B3LYP-D	0.95	0.86	0.41	0.40	0.36	0.37	0.38	0.38
M06-2X <sup>a</sup>	0.38	0.48	0.26	0.50	0.62	0.95	0.51	0.82
$\omega$ B97X-D		0.64		0.49		0.97		0.82
B3LYP-XDM		0.53		0.23		0.42		0.37
<b>6-311+G(3df,2p)</b>								
XYG3		0.58		0.09		0.15		0.13
B3LYP-D	0.67	0.58	0.26	0.29	0.26	0.15	0.26	0.20
<b>aug-cc-pVTZ</b>								
B97		2.43		2.86		4.31		3.87
BP86		1.20		2.46		3.87		3.44
B3LYP		1.13		2.39		3.85		3.40
B970		1.16		1.92		3.13		2.76
PBE		0.42		1.54		2.60		2.28
PBE0		0.39		1.47		2.64		2.29
B2PLYP		0.53		1.13		1.86		1.64
XYG3		0.46		0.64		1.29		1.09
M05-2X <sup>a</sup>	0.46	0.63	0.38	0.24	0.88	0.23	0.73	0.23
PBE0-D	1.04	1.12	0.14	0.24	0.53	0.21	0.41	0.22
PBE-D	0.95	0.75	0.19	0.24	0.15	0.26	0.17	0.25
B970-D	0.25		0.20		0.48		0.40	
BP86-D	0.74	0.81	0.41	0.37	0.39	0.56	0.39	0.50
B2PLYP-D	0.40	0.36	0.19	0.14	0.22	0.17	0.21	0.16
B97-D	0.58	0.37	0.14	0.21	0.16	0.07	0.15	0.11
B3LYP-D	0.65	0.55	0.20	0.20	0.20	0.13	0.20	0.15
M06-2X <sup>a</sup>	0.30	0.35	0.31	0.16	0.31	0.22	0.31	0.20
$\omega$ B97X-D		0.37		0.18		0.33		0.29
B3LYP-XDM		0.25		0.18		0.09		0.12

<sup>a</sup>Column labeled -D2 (-D3) contains figure for M0N-2X (M0N-2X-D3) functional variant.

are not available. An earlier study by Karton and *et al.*<sup>100</sup> similarly showed a -D2 correction to M06-2X with the pc-2 basis ( $\sim$ cc-pVTZ) to be ineffective, with an optimal  $s_6^{\text{M06-2X}}$  of 0.06. Combined with the lack of dominance by either M0N-2X or M0N-2X-D3 across basis sets, the main discussion will focus on the performance of the uncorrected parent functionals, though relevant statistics for both variants are given in Tables III–VI. In the current work, superior overall performance (M06-2X-D3 versus M06-2X) with aTZ follows from substantial MAD reduction (accompanied by better alignment of

errors among test sets) within the MX and DD subsets so that, in the context of the ordinals of Sec. III C 3, M05-2X-D3 and M06-2X-D3 fall on either side of M06-2X, which is a slight promotion for the latter and a great improvement for the former. Despite its mixed success, the -D3 term is found to ameliorate many of the qualitative problems discussed later for M05-2X and M06-2X, particularly the occurrence of simultaneous underbinding and overbinding deviations for a model chemistry and system-size dependent error trends. These improvements at both the D $\zeta$  and T $\zeta$  levels suggest that a “dash

TABLE V. Mean absolute deviation of the interaction energy (kcal/mol) for the HSG test set with noncounterpoise-corrected DFT methods.

Method and basis set	Hydrogen bonded		Mixed influence		Dispersion bound		HSG	
	-D2	-D3	-D2	-D3	-D2	-D3	-D2	-D3
<b>aug-cc-pVDZ</b>								
B97	4.26		2.51		2.65		2.92	
BP86	2.66		2.44		2.24		2.38	
B3LYP	2.17		1.88		1.90		1.95	
B970	2.13		1.40		1.41		1.55	
PBE	1.13		1.08		0.97		1.03	
PBE0	0.98		1.21		1.09		1.10	
B2PLYP	0.74		0.58		0.52		0.58	
XYG3	0.85		0.41		0.50		0.54	
M05-2X <sup>a</sup>	0.40	1.03	0.15	0.77	0.17	0.90	0.21	0.89
PBE0-D	0.76	1.03	0.23	0.72	0.63	0.89	0.54	0.87
PBE-D	1.05	0.86	0.71	0.86	1.18	1.01	1.02	0.94
B970-D	0.46		0.40		0.74		0.59	
BP86-D	0.38	0.69	0.33	0.58	0.77	0.94	0.57	0.79
B2PLYP-D	0.86	0.78	0.73	0.77	1.06	0.90	0.93	0.84
B97-D	0.80	0.58	0.47	0.47	0.93	0.70	0.78	0.61
B3LYP-D	0.87	0.74	0.63	0.76	1.11	0.86	0.93	0.81
M06-2X <sup>a</sup>	0.61	0.95	0.20	0.45	0.41	0.80	0.39	0.73
$\omega$ B97X-D	0.49		0.49		0.97		0.74	
B3LYP-XDM	0.44		0.39		0.57		0.49	
<b>6-311+G(3df,2p)</b>								
XYG3	0.53		0.19		0.23		0.27	
B3LYP-D	0.57	0.45	0.25	0.38	0.52	0.27	0.45	0.33
<b>aug-cc-pVTZ</b>								
B97	4.62		2.92		3.22		3.40	
BP86	3.05		2.85		2.87		2.90	
B3LYP	2.63		2.31		2.52		2.48	
B970	2.56		1.79		1.99		2.04	
PBE	1.57		1.49		1.55		1.54	
PBE0	1.41		1.57		1.62		1.56	
B2PLYP	0.84		0.95		1.08		1.00	
XYG3	0.37		0.23		0.28		0.28	
M05-2X <sup>a</sup>	0.24	0.76	0.42	0.40	0.58	0.21	0.47	0.37
PBE0-D	0.36	0.60	0.19	0.36	0.20	0.36	0.23	0.40
PBE-D	0.60	0.41	0.31	0.45	0.60	0.43	0.52	0.43
B970-D	0.50		0.13		0.18		0.23	
BP86-D	0.32	0.30	0.52	0.30	0.30	0.32	0.37	0.31
B2PLYP-D	0.75	0.68	0.37	0.40	0.50	0.34	0.51	0.42
B97-D	0.99	0.68	0.23	0.20	0.36	0.15	0.45	0.26
B3LYP-D	0.53	0.45	0.20	0.33	0.49	0.24	0.41	0.31
M06-2X <sup>a</sup>	0.43	0.62	0.43	0.13	0.36	0.13	0.39	0.22
$\omega$ B97X-D	0.29		0.18		0.41		0.32	
B3LYP-XDM	0.32		0.08		0.11		0.14	

<sup>a</sup>Column labeled -D2 (-D3) contains figure for M0N-2X (M0N-2X-D3) functional variant.

D'' correction is indeed beneficial to the M0N-2X functional class.

The DFT-D3 fitting parameters suggested by Grimme *et al.* (D3-def) were obtained using a quadruple- $\zeta$  basis set, essentially beyond the interference of BSSE. Alternate values for triple- $\zeta$  calculations (D3-alt) are encoded into the DFTD3 program,<sup>66</sup> though the developers found mixed results<sup>19</sup> for these in preliminary testing. The application of alternate parameters to the test sets in this work yields nearly universal improvement for the aug-cc-pVDZ basis set, with average reduction in error ( $\text{MAD}^{\text{DFT-D3-alt}} - \text{MAD}^{\text{DFT-D3-def}}$ ) of  $-0.13$ ,  $-0.11$ ,  $-0.17$ , and  $-0.13$  kcal/mol for HB, MX, DD, and

overall subsets, respectively. The HB category with B97-D3 presents a clear exception by underperforming with D3-alt for both aDZ and aTZ basis sets, reflecting, perhaps, a disruption to closely bound systems by the heavily weighted -D correction in this functional. Apart from that case, the aug-cc-pVTZ patterns are inconclusive, with HB and MX favoring D3-alt and DD favoring D3-def for a net nil effect among overall errors. Although the alternate TZ values are nominally more appropriate for this work, which employs only basis sets of less than quadruple- $\zeta$  quality, all further DFT-D3 results follow from the default parameters to better assess the standard form of this new and promising method. A table with DFT-



TABLE VI. Mean absolute deviation of the interaction energy (kcal/mol) for the JSCH test set with noncounterpoise-corrected DFT methods.

Method and basis set	Hydrogen bonded		Interstrand		Stacked		JSCH	
	-D2	-D3	-D2	-D3	-D2	-D3	-D2	-D3
<b>aug-cc-pVDZ</b>								
B97	5.88		2.09		10.79		7.04	
BP86	2.99		2.19		9.06		5.43	
B3LYP	2.83		1.89		8.59		5.10	
B970	3.07		1.41		7.41		4.53	
PBE	1.44		1.14		6.42		3.53	
PBE0	1.18		1.26		6.12		3.35	
B2PLYP	1.06		0.73		3.18		1.90	
XYG3	0.61		0.44		1.33		0.88	
M05-2X <sup>a</sup>	0.83	0.48	0.54	0.33	1.19	0.75	0.91	0.56
PBE0-D	1.55	1.87	0.14	0.40	0.76	0.37	0.84	0.84
PBE-D	1.94	1.55	0.32	0.53	0.47	0.52	0.88	0.84
B970-D	0.53		0.13		0.72		0.51	
BP86-D	1.68	2.12	0.25	0.30	0.64	1.49	0.86	1.37
B2PLYP-D	1.45	1.32	0.34	0.37	1.73	1.41	1.29	1.11
B97-D	0.80	0.61	0.35	0.19	0.56	0.54	0.58	0.47
B3LYP-D	1.84	1.58	0.21	0.28	0.83	0.50	0.98	0.78
M06-2X <sup>a</sup>	0.61	0.41	0.41	0.19	0.55	1.19	0.53	0.69
$\omega$ B97X-D	1.06		0.28		1.03		0.84	
B3LYP-XDM	0.71		0.11		0.26		0.36	
<b>6-311+G(3df,2p)</b>								
XYG3	0.75		0.17		0.24		0.38	
B3LYP-D	0.98	0.72	0.16	0.18	0.47	0.33	0.55	0.41
<b>aug-cc-pVTZ</b>								
B97	6.50		2.27		11.53		7.60	
BP86	3.70		2.39		9.87		6.05	
B3LYP	3.60		2.11		9.32		5.71	
B970	3.72		1.64		8.28		5.17	
PBE	2.14		1.36		7.30		4.19	
PBE0	1.75		1.48		7.05		3.99	
B2PLYP	1.69		1.05		4.58		2.78	
XYG3	0.72		0.51		2.14		1.28	
M05-2X <sup>a</sup>	1.15	0.53	0.78	0.21	2.22	0.43	1.52	0.40
PBE0-D	0.95	1.25	0.33	0.22	1.69	0.80	1.11	0.79
PBE-D	1.21	0.87	0.14	0.33	0.61	1.00	0.67	0.79
B970-D	0.55		0.21		1.58		0.91	
BP86-D	1.02	1.40	0.37	0.22	0.66	0.76	0.69	0.81
B2PLYP-D	0.96	0.83	0.13	0.12	0.42	0.21	0.51	0.38
B97-D	1.06	0.78	0.20	0.18	0.48	1.13	0.59	0.78
B3LYP-D	1.13	0.84	0.15	0.15	0.34	0.42	0.53	0.48
M06-2X <sup>a</sup>	1.06	0.62	0.65	0.19	0.81	0.38	0.85	0.41
$\omega$ B97X-D	0.49		0.17		0.29		0.32	
B3LYP-XDM	0.37		0.21		0.59		0.42	

<sup>a</sup>Column labeled -D2 (-D3) contains figure for M0N-2X (M0N-2X-D3) functional variant.

D2, DFT-D3-def, and DFT-D3-alt statistics is available in the repository of supplementary materials.<sup>85</sup>

Of the several variants that exist for a functional (i.e., DFT, DFT-D2, and DFT-D3), this section has extracted guidelines on which one is the most widely accurate. In keeping with these findings, the comparisons of functional performance that follow in Secs. III B and III C employ the M05-2X, M06-2X, B970-D2, BP86-D2, B3LYP-D3, B97-D3, PBE0-D2, PBE-D3, and B2PLYP-D3 forms, where the last three are replaced by PBE0, PBE, and B2PLYP

when referring specifically to hydrogen-bonded systems in the double- $\zeta$  regime.

## B. Comparison of methods with double- $\zeta$ basis sets

The primary results for DFT-D, specialized functionals, and XDM with the aug-cc-pVDZ basis are summarized through MAD statistics in the upper portion of Table III for S22, Table IV for NBC10 and HBC6, Table V for HSG, and Table VI for JSCH, as well as represented graphically in Fig.

TABLE VII. Comparison of noncounterpoise- and CP-corrected results for six hierarchical double- $\zeta$  and triple- $\zeta$  basis sets (additional doubly augmented sets available for B3LYP-D3). The mean absolute deviation of the interaction energy (kcal/mol) is presented for the S22 test set and subsets.

Method and basis set	non-CP-corrected				CP-corrected			
	H-bonded	Mixed	Dispersion	S22	H-bonded	Mixed	Dispersion	S22
<b>M05-2X</b>								
cc-pVDZ	2.77	0.51	0.49	1.22	1.16	0.59	1.54	1.12
heavy-aug-cc-pVDZ	0.58	0.31	0.26	0.38	0.92	0.48	1.03	0.82
aug-cc-pVDZ	0.45	0.43	0.21	0.36	0.90	0.49	1.02	0.81
cc-pVTZ	0.92	0.38	0.70	0.67	0.79	0.49	1.24	0.86
heavy-aug-cc-pVTZ	0.63	0.40	0.89	0.65	0.73	0.49	1.17	0.81
aug-cc-pVTZ	0.61	0.40	0.89	0.64	0.72	0.49	1.18	0.81
<b>B970-D2</b>								
cc-pVDZ	3.55	0.92	0.59	1.64	0.57	0.29	1.15	0.69
heavy-aug-cc-pVDZ	0.19	0.41	0.27	0.29	0.51	0.33	0.90	0.60
aug-cc-pVDZ	0.25	0.57	0.24	0.35	0.50	0.32	0.93	0.60
cc-pVTZ	0.98	0.44	0.41	0.60	0.42	0.31	0.97	0.58
heavy-aug-cc-pVTZ	0.27	0.30	0.73	0.44	0.41	0.31	0.89	0.56
aug-cc-pVTZ	0.23	0.28	0.72	0.43	0.41	0.31	0.89	0.55
<b>B3LYP-D3</b>								
cc-pVDZ	4.72	1.28	1.33	2.39	0.27	0.14	0.50	0.31
heavy-aug-cc-pVDZ	0.86	0.46	0.49	0.60	0.33	0.14	0.31	0.26
aug-cc-pVDZ	1.06	0.67	0.66	0.79	0.34	0.13	0.34	0.27
d-aug-cc-pVDZ	1.19	0.78	0.80	0.92	0.34	0.13	0.34	0.27
cc-pVTZ	1.83	0.70	0.47	0.98	0.45	0.16	0.33	0.32
heavy-aug-cc-pVTZ	0.57	0.22	0.21	0.33	0.42	0.14	0.29	0.29
aug-cc-pVTZ	0.63	0.26	0.22	0.36	0.43	0.15	0.29	0.29
d-aug-cc-pVTZ	0.65	0.27	0.23	0.37	0.43	0.15	0.29	0.29
<b>M06-2X</b>								
cc-pVDZ	2.66	0.55	1.08	1.42	1.07	0.41	0.49	0.65
heavy-aug-cc-pVDZ	0.50	0.27	0.88	0.57	0.82	0.30	0.18	0.42
aug-cc-pVDZ	0.38	0.50	1.14	0.69	0.78	0.29	0.19	0.41
cc-pVTZ	0.83	0.36	0.41	0.53	0.76	0.36	0.28	0.46
heavy-aug-cc-pVTZ	0.59	0.28	0.22	0.36	0.69	0.34	0.22	0.41
aug-cc-pVTZ	0.54	0.27	0.22	0.34	0.67	0.35	0.21	0.40
<b><math>\omega</math>B97X-D</b>								
cc-pVDZ	3.61	0.94	1.42	1.97	0.33	0.13	0.27	0.24
heavy-aug-cc-pVDZ	0.49	0.55	1.04	0.71	0.18	0.18	0.33	0.23
aug-cc-pVDZ	0.64	0.74	1.23	0.89	0.19	0.12	0.36	0.23
cc-pVTZ	1.21	0.53	0.66	0.79	0.23	0.13	0.16	0.17
heavy-aug-cc-pVTZ	0.22	0.25	0.41	0.30	0.21	0.15	0.17	0.18
aug-cc-pVTZ	0.27	0.30	0.42	0.33	0.21	0.15	0.18	0.18

2. Assessments of each method for purely hydrogen-bonded systems, purely dispersion-bound systems and generic NCI systems follow.

### 1. Hydrogen-bonded systems

In accordance with the analysis in Sec. III A and in order of increasing performance, the BP86-D2, B3LYP-D3, PBE, PBE0,  $\omega$ B97X-D, XYG3, B2PLYP, M05-2X, M06-2X, B97-D3, and B970-D2 functionals are compared for hydrogen-bonded complexes at the double- $\zeta$  level of theory. The least effective of these, BP86-D2 and B3LYP-D3, are significantly overbound, with the error for most individual systems lying beyond  $-1$  kcal/mol. To a lesser degree, the PBE and PBE0 functionals provide similar deviation patterns on the underbound side. Of the previous, BP86-D2 and PBE0 have pervasive “parallelism” trouble replicating the shape of potential energy curves in HBC6, as evinced by the widely

scattered errors for this set in Fig. 2. The  $\omega$ B97X-D approach has two MAD markers outside (overbound) of 1 kcal/mol for HBC6 and JSCH but nevertheless exhibits a satisfactory precision compared to the preceding techniques.

Six methods perform creditably for HB/D $\zeta$  systems by possessing all test set MAD markers approaching or within 1 kcal/mol. The XYG3 approach tends toward overbinding, generating its least satisfying results with the HBC6 set, for which, in common with M05-2X and M06-2X, it displays difficulty following the curve contour for the systems with the most charge-transfer character, namely, those complexed with FaOO. B2PLYP, in contrast, is underbound (in common with most uncorrected functionals) but elicits highly consistent individual errors, generally within  $+0.3$  to  $+1.0$  kcal/mol. Its largest inaccuracies are with the JSCH set, for which the reference energies themselves are less reliably established. Both M05-2X and M06-2X underbind (JSCH) and overbind (HBC6) certain test sets, yet are among the best-behaved for S22 and HSG. The final three of the introductory list are the

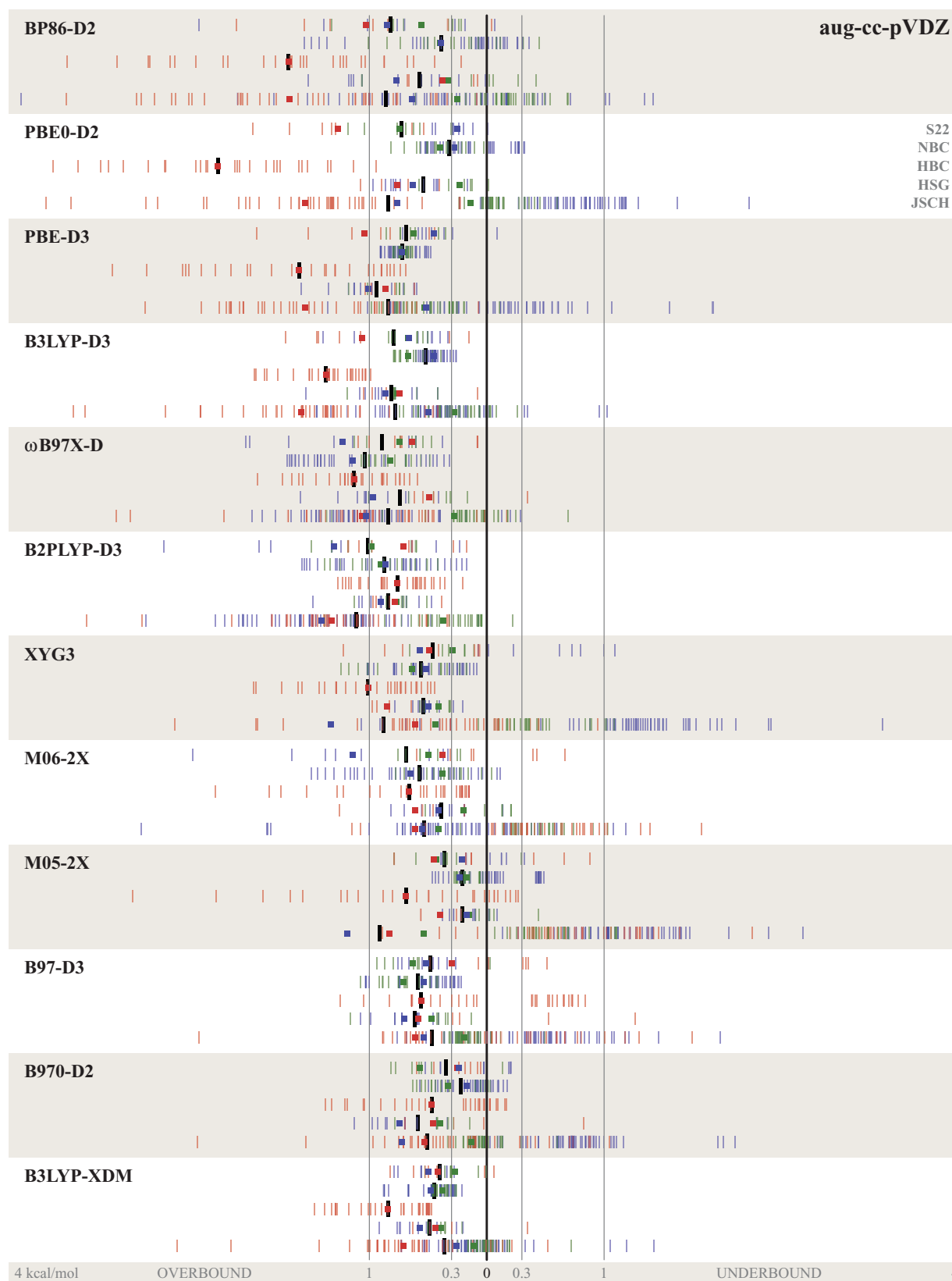


FIG. 2. Double- $\zeta$  functional performance. For each DFT technique considered, individual member errors and MAD summary statistics are plotted for five test sets at the non-CP-corrected aug-cc-pVDZ level of theory. Consult Sec. II D in the text for a concise guide to the plotting mode used here.

premier performers, with all MAD less than 0.65 kcal/mol. In these functionals, individual member errors are rarely outside  $\pm 1$  kcal/mol and are generally evenly distributed among overbound and underbound interaction energies. Although not approaching the desirable accuracy of  $\text{MAD} < 0.3$  kcal/mol, M06-2X, B97-D3, and B970-D2 are the most recommended. Despite the varying levels of success for hydrogen-bonded systems that emerge from an examination of absolute error, the large binding energies characteristic of this class make relative deviation criteria undemanding, and all methods considered here meet the strict ( $\text{RQ}_1 < 10\%$ ) cutoff. The underlying functionals (with the exception of B97) generally satisfy the looser (20%) standard, in keeping with DFT's satisfactory reputation for electrostatically dominated complexes, though van der Waals interactions have been found to be of importance even for these.<sup>60</sup>

## 2. Dispersion-bound systems

Complexes bound by dispersion interactions are treated erratically at the double- $\zeta$  level. In order of increasing success are the functionals B2PLYP-D3,  $\omega$ B97X-D, M06-2X, BP86-D2, PBE-D3, XYG3, B3LYP-D3, B97-D3, PBE0-D2, B970-D2, and M05-2X. Unlike the HB subgroup, where membership is well-defined and the mean interaction energies are comparable across the test sets, the analog to the DD category in JSCH is defined by monomer configuration ("stacked" nucleobases that, on the basis of SAPT calculations, would be partitioned into both MX and DD classes) and consequently has different composition and deeper average IE than the others (*cf.* Table II). Coupled with the less accurately fixed reference values of JSCH, this test set is somewhat discounted in evaluating functional performance for purely dispersion-bound systems. The worst methods are B2PLYP-D3 and  $\omega$ B97X-D which are systematically overbound having absolute deviations centered around  $-1$  kcal/mol. Although not shifted to such a severe degree, M06-2X and BP86-D2 display a very broad scattering of individual errors. These latter two approaches also exhibit acute parallelism difficulties tracing the relevant PEC in NBC10; this has been noted previously for the former technique.<sup>58</sup>

The seven acceptable methods for DD/D $\zeta$  possess individual deviations within  $\pm 1$  kcal/mol for almost all members of non-JSCH test sets. PBE-D3 is the first functional of the list with tight error variance, though its MAD for HSG lies at  $-1.0$  kcal/mol, the same magnitude as IE quantities for this subset. The dispersion-dominated members of HSG in general prove difficult for many DFT methods, being of nonequilibrium and sometimes unbound configuration. At the aDZ level, only M05-2X and M06-2X achieve tolerable success, a testament to their optimization for midrange interactions. Following PBE-D is XYG3 which, for S22, has the distinction of being the only functional to underbind a test set (excluding JSCH), whereas for NBC10 and HSG it exhibits the usual overbinding. B3LYP-D3 and B97-D3 are the most consistent techniques for the four test sets, with MAD from 0.45–0.86 kcal/mol for the former and from 0.52–0.70 for the latter, though both trend over-

bound, with few non-JSCH members approaching the zero-line. Starting with PBE0-D2 and B970-D2, the S22 and NBC10 sets achieve an average accuracy better than 0.3 kcal/mol, though these methods have rather worse performance for the other test sets. The final functional, M05-2X, is considerably underbound for JSCH, but performs beautifully for the remaining test sets, with errors nicely centered around the zero-line and essentially all within  $\pm 0.5$  kcal/mol. In contrast to HB, MA%BD for the DD subset presents a strict assessment due to its relatively weaker interaction energies, and only PBE0-D2 and B970-D2 pass even the loose criterion for test sets excluding HSG.

## 3. Comprehensive double- $\zeta$ results

Evaluating DFT techniques to obtain a balanced performance for all noncovalently interacting systems with the aug-cc-pVDZ basis set yields the following ordering, toward the most successful: BP86-D2, PBE0-D2, PBE-D3, B3LYP-D3,  $\omega$ B97X-D, B2PLYP-D3, XYG3, M06-2X, M05-2X, B97-D3, and B970-D2. In Sec. III B 1, the first four methods were rejected for hydrogen-bonding systems due to their extreme overbinding which places nearly all individual deviations outwards of  $-1$  kcal/mol. (Notice that the (overbinding) DFT-D variants of PBE and PBE0 are considered here rather than the (underbinding) underlying functionals.) Especially for PBE0-D2 and PBE-D3, these approaches exhibit "banded" results where HB errors fall at more negative values of the range, DD errors fall at more positive values, and MX errors lie in the middle (some degree of sorting is expected, given the classes' relative IE). An implication of the finding that such functionals have disparate performance for different types of NCI is the risk that any good results for a compound system may only be due to cancellation of error. Also among the unfavored, the  $\omega$ B97X-D and B2PLYP-D3 approaches were the lowest ranked for dispersion dominated systems, though these deviations are less dramatic than the HB failures of the previous techniques, and produce overall MAD markers around 0.9 kcal/mol.

Five D $\zeta$  model chemistries are found to treat NCI systems in a satisfactory manner, notwithstanding earlier recommendations that DFT-D approaches should be coupled with T $\zeta$  or larger bases.<sup>54,56</sup> Among these, the all-inclusive test set errors are roughly consistent, with averages descending from 0.58 to 0.41 kcal/mol through the list (*cf.* Table I), but patterns among the individual member deviations introduce some reservations even for the favored methods. The XYG3 functional has overall MAD markers (and almost all subset quantities) within 1 kcal/mol, yet for JSCH and the very-accurately benchmarked S22, it exhibits both banding and a wide error span. Although the only technique not recommended in both HB and DD subcategories (the latter due to problems of overbinding and parallelism), M06-2X rates a second-class approval overall, despite its broad error range and slight banding for S22 and JSCH. For the test sets at which M05-2X excels, it has essentially the best statistics for the methods considered. However, the functional overbinds HBC6 (recall the parallelism difficulties that affect this and the two preceding techniques) and modestly underbinds JSCH, leading to



a large uncertainty in the position and direction of deviation for any unbenchmarked interaction energy. The final two approaches, B97-D3 and B970-D2, have clean and consistent performances. Their test set MADs are within 0.47–0.61 and 0.22–0.59 kcal/mol, respectively, no subset marker exceeds 0.75 kcal/mol, no test sets have an abnormally large variance, and almost all individual member errors are within 1 kcal/mol of the zero-line. These are the most unreservedly recommended; nevertheless, both have deviations which are in magnitude and spread larger than the best  $T\zeta$  model chemistries.

Considering Hobza's  $RQ_1$  requirement applied to overall statistics, three approaches, M05-2X, B970-D2, and PBE0-D2, satisfy the criterion for four out of five test sets (HSG or JSCH excluded). The presence of the last functional serves as a reminder of how poor HB performance is not filtered out by this metric. Another technique that is positioned among the best-ranked is B3LYP-XDM, which confines most individual deviations within  $-1.0$  to  $+0.3$  and presents a low variance in error. These aDZ XDM results follow from parameters optimized for the aTZ basis set.

### C. Comparison of methods with triple- $\zeta$ basis sets

The primary results for DFT-D, specialized functionals, and XDM with the aug-cc-pVTZ and 6-311+G(3df,2p) bases are summarized through MAD statistics in the lower portion of Table III for S22, Table IV for NBC10 and HBC6, Table V for HSG, and Table VI for JSCH, as well as represented graphically in Fig. 3. Assessments of each method for purely hydrogen-bonded systems, purely dispersion-bound systems, and generic NCI systems follow.

#### 1. Hydrogen-bonded systems

A comparison of DFT techniques for HB/ $T\zeta$  complexes yields the following ordering by increasing performance: PBE0-D2, BP86-D2, M05-2X, PBE-D3, XYG3/aTZ, M06-2X, XYG3/pp2, B3LYP-D3, B97-D3, B2PLYP-D3,  $\omega$ B97X-D, and B970-D2. The best and worst error statistics for hydrogen-bonded systems are essentially constant between aDZ and aTZ levels of theory; therefore, appreciable change in ordinal for a given functional between this and the HB/D $\zeta$  list reflects a corresponding trend in basis set dependence, rather than any wholesale improvement in describing HB interactions by the larger basis. The least promising methods, PBE0-D2 and BP86-D2, are severely overbound, with very wide spreads in individual deviation and averages centered at about  $-1$  kcal/mol. The M05-2X functional, which was one of the preferred for aDZ, is notably inferior here, primarily because of greater error variance for every test set, as well as mixed underbinding and overbinding tendencies. In contrast, PBE-D3 is firmly overbound, presenting a more compact range of deviation than the preceding methods (or than PBE/aDZ). While XYG3 MAD markers for the 6-311+G(3df,2p) basis employed in its development are somewhat inferior to aDZ, both bases achieve superior error precision compared to aug-cc-pVTZ, which in particular is erratic for the hydrogen-bonded curves. This shortcoming is due

in part to the same parallelism difficulties tracing HBC6 PEC by BP86-D2, PBE0-D2, XYG3 (Dunning basis sets only), M05-2X, and M06-2X discussed in Sec. III B 1.

Among the functionals deemed acceptable, all have test set MAD positions approaching or within 1 kcal/mol. Like B2PLYP in the HB/D $\zeta$  case, M06-2X exceeds this key value only for JSCH, which is less reliably benchmarked than the other sets. The dependability of this technique is also hampered by its widely spread individual errors, even within S22. The B3LYP-D3 functional exhibits significantly improved performance with both triple- $\zeta$  basis sets considered in terms of both mean and distribution of error, while trending reliably overbound. Occupying a similar range of average deviations is B97-D3 which, despite slightly worse MAD values compared to aDZ, remains a technique with few member errors outside  $\pm 1$  kcal/mol. At comparable MAD values to B3LYP-D3 and B97-D3, B2PLYP-D3 has the smallest spread in errors of the methods examined, but because it tends to overbind, there are actually very few deviations for H-bonded complexes within the desirable  $\pm 0.3$  kcal/mol window. The  $\omega$ B97X-D technique is clearly better attuned to triple- $\zeta$  basis sets, achieving lower MAD for all test sets and (except for HBC6) nearly the absolute lowest MAD among all model chemistries. The B970-D2 functional exceeds this performance only in providing deviations that are very cleanly centered over the zero-line and most likely to lie within  $\pm 0.3$  kcal/mol. The recommended final four functionals, B97-D3, B2PLYP-D3,  $\omega$ B97X-D, and B970-D2, display both accuracy and precision in interaction energy errors, yet it is the latter quantity by which aTZ is slightly superior to aDZ, rather than the former (for which aDZ is arguably ascendant). Although ever dependent on particular circumstances, there is no clear advantage to triple- $\zeta$  basis sets for purely hydrogen-bonding systems. Like the double- $\zeta$  case, in terms of relative error all functionals meet the strict  $RQ_1$  criterion, with the best methods falling below 6%.

#### 2. Dispersion-bound systems

Functionals in order of increasing success for dispersion-dominated systems with triple- $\zeta$  basis sets are XYG3/aTZ, M05-2X, B970-D2, PBE0-D2, PBE-D3, BP86-D2, M06-2X, B97-D3,  $\omega$ B97X-D, B2PLYP-D3, B3LYP-D3, and XYG3/pp2. A conspicuous characteristic of this list is that the five worst-performing functionals on it were highly rated with aug-cc-pVDZ; the converse is true for unacceptable DD/aDZ model chemistries. The XYG3 and M05-2X approaches perform very poorly with aug-cc-pVTZ, possessing error spreads far greater than those for any functional with aDZ. B970-D2 and PBE0-D2 are better behaved, with most individual (non-JSCH) deviations within  $-0.3$  to  $+1.0$  kcal/mol, but still slip from their aug-cc-pVDZ performance, rather than capitalizing on the triple- $\zeta$  basis set as do others. The preceding four methods show an atypical, consistent trend toward underbinding, and more troubling is their difficulty in tracing the PEC for complexes in NBC10 with sandwich orientations. A final technique, PBE-D3

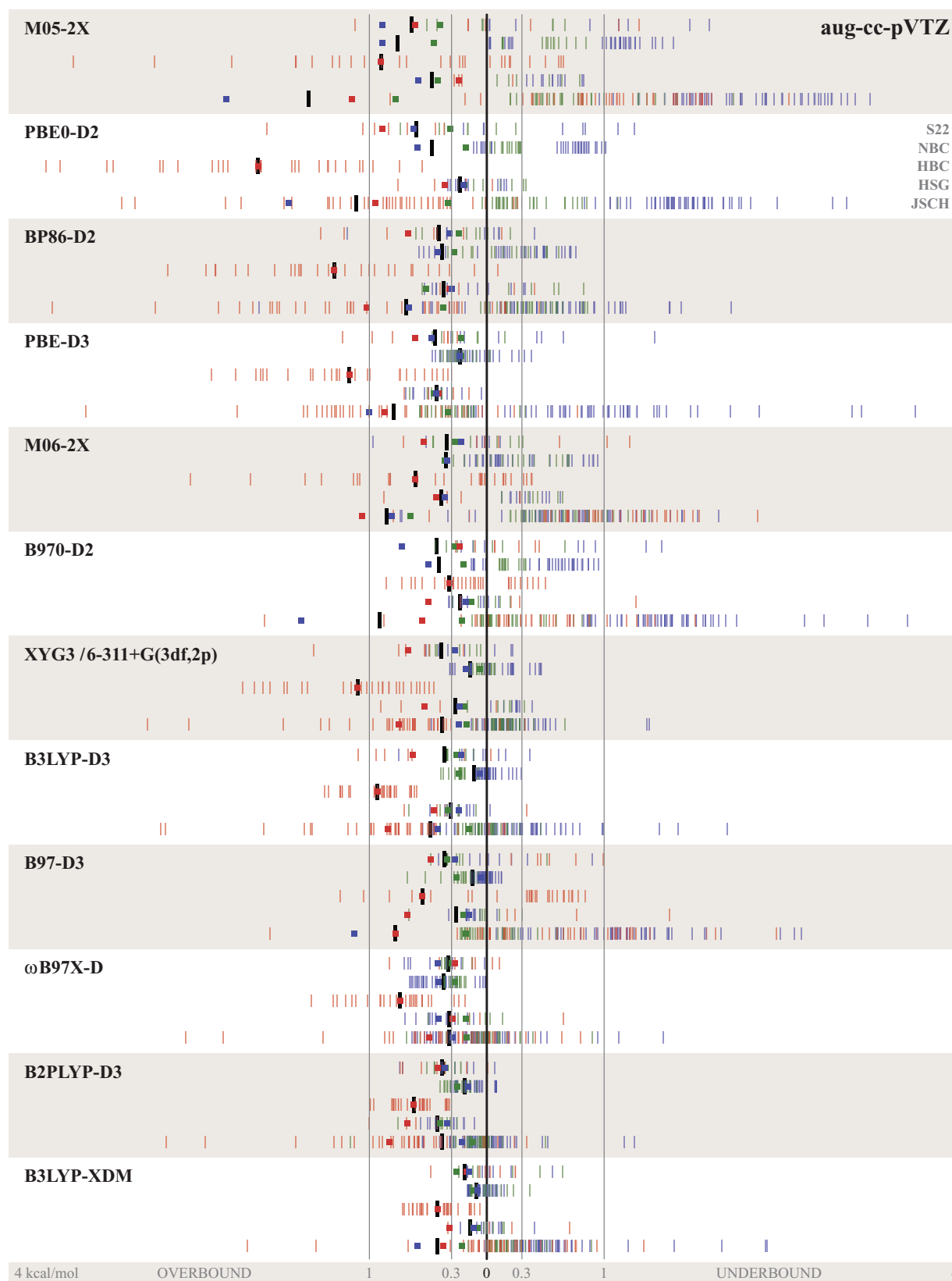


FIG. 3. Triple- $\zeta$  functional performance. For each DFT technique considered, individual member errors and MAD summary statistics are plotted for five test sets at the non-CP-corrected aug-cc-pVTZ level of theory. Consult Sec. II D in the text for a concise guide to the plotting mode used here.

is borderline acceptable, yet remains demoted by its high variance for the S22 and JSCH test sets.

Each of the seven successful DFT methods for DD/T $\zeta$  produces interaction energy error distributions that are shifted positively (toward underbinding) compared to aug-cc-pVDZ, with the resultant (non-JSCH) MAD markers entirely below 0.42 kcal/mol. The first of these, BP86-D2 and M06-2X, have a larger error variance than the succeeding functionals and exhibit general trouble tracking the minimum for NBC10 curves (the former more severely). For B97-D3, whose test set MAD were all aligned around 0.55 kcal/mol with aDZ, transition to the larger basis has yielded exemplary deviations ( $<0.3$  kcal/mol), except for JSCH which is overcorrected.  $\omega$ B97X-D and B2PLYP-D3 give mean errors centered around 0.3 kcal/mol as well as very tight error clustering that is much improved compared to aDZ. Meanwhile, B3LYP-D3 maintains its double- $\zeta$  MAD patterns for both triple- $\zeta$  basis sets, while shifting IE deviations significantly towards the zero-line. In association with its development basis, XYG3 affords outstanding statistics, as observed previously for NBC10,<sup>59</sup> with almost all individual deviations between  $\pm 0.3$  kcal/mol. These last five, B97-D3,  $\omega$ B97X-D, B2PLYP-D3, B3LYP-D3, and XYG3/6-311+G(3df,2p) are the most highly recommended, of which all nearly pass the lenient relative error requirement for three or all of the test sets.

### 3. Comprehensive triple- $\zeta$ results

In order of increasing success for generic NCI complexes with triple- $\zeta$  basis sets are the DFT techniques XYG3/aTZ, M05-2X, PBE0-D2, BP86-D2, PBE-D3, M06-2X, B970-D2, XYG3/pp2, B3LYP-D3, B97-D3,  $\omega$ B97X-D, and B2PLYP-D3. The plainly unsuccessful first five of these are strongly banded, supporting many hydrogen-bonding error bars beyond  $-1$  kcal/mol and many dispersion-bound error bars beyond  $+1$  kcal/mol. Similarly to the aDZ case, the unsatisfactory treatments occupy the worst positions on the HB and DD lists, and their poor performance is amplified by the former being overbound and the latter being underbound, such that the full error range spans 6–8 kcal/mol. The caveats on parallelism introduced in earlier subsections for all of these functionals but PBE-D3 should also be considered. While all five techniques exhibit a performance decline for aTZ compared to aDZ, only for those functionals that excelled with the smaller basis set, XYG3 and M05-2X, is this due to larger MAD quantities. For the others, mean deviations actually improve at the triple- $\zeta$  level, but this is counteracted by broader error spread.

The seven remaining DFT methods classified as successful for TT/T $\zeta$  produce average test set MAD values of 0.33–0.44 kcal/mol (*cf.* Table I), which are significantly better than the functionals in comparable positions for the aug-cc-pVDZ basis set. The origin of this shift is not due to hydrogen-bonded systems that actually show no change ( $0.53 \leftarrow 0.51$  kcal/mol) but is due to the dispersion-bound systems that are appreciably improved ( $0.38 \leftarrow 0.56$  kcal/mol), as discussed more fully in Sec. III D. The first, M06-2X, continues its aDZ difficulty with simultaneously overbinding HBC6 (for which

it also exhibits parallelism trouble) and trending toward underbinding others, but it otherwise constricts error ranges so that almost all S22, NBC10, and HSG members are within  $\pm 1$  kcal/mol. Compared to aDZ, B970-D2 distinctly shifts its error distribution roughly a half kcal/mol toward underbinding, yielding better overall MAD values for all but the JSCH test set. The difficulty with JSCH results from a tendency of the method to compute greater magnitude IE errors for dispersion-dominated complexes as system size increases, which was reflected in B970-D2 not being a recommended functional in the DD class. Next are XYG3 in conjunction with its development Pople-style basis and B3LYP-D3/aTZ, which generate similar levels of error overall and among the subsets. Unfortunately both also exhibit banding for almost all test sets; however, since the actual deviations are small (generally within  $-1.0$  to  $+0.3$  kcal/mol for non-JSCH) and of very low variance, this is not considered a serious weakness. Though the best performer for dispersion-bound systems, B3LYP-D3 is systematically, albeit modestly, overbound for the HB class, hence the banding. Similarly to B970-D2, B97-D3 interaction energy deviations have shifted in a positive direction with accordant effect on overall MAD values. Contrastingly, B97-D3 is the most clearly stable functional with respect to system size, showing the same error distribution (in MA%D) for large as well as small complexes. The  $\omega$ B97X-D and B2PLYP-D3 functionals, while computationally expensive, perform outstandingly, having almost all individual deviations within  $-1.0$  to  $0.0$  kcal/mol for non-JSCH and  $-1.0$  to  $+1.0$  kcal/mol for JSCH, while maintaining a low variance for each test set.  $\omega$ B97X-D excels particularly in resisting banding. That is, its subgroup MAD markers for each test set are remarkably similar, which, considering the differences in typical IE magnitude among the subsets (*cf.* Table II) is a testament to its success for hydrogen-bonded systems. B2PLYP-D3 maintains subset averages for HB that are somewhat greater than those of MX and DD and presents slightly more compact error distributions within test sets, especially for HBC6, but is otherwise very comparable to  $\omega$ B97X-D. The XYG3/6-311+G(3df,2p), B3LYP-D3, B97-D3,  $\omega$ B97X-D, and B2PLYP-D3 functionals, particularly the last two, are recommended to obtain the best results for an arbitrary NCI complex.

Overall success, as measured by relative error, is achieved by B97-D3 and XYG3/6-311+G(3df,2p), both of which pass MA%BD  $< 20\%$  for all five test sets. Joining their company is the B3LYP-XDM method, whose performance is competitive with the best T $\zeta$  approaches at considerably less cost. Apart from these, the other highly recommended functionals, B3LYP-D3,  $\omega$ B97X-D, and B2PLYP-D3 also fulfill the RQ<sub>1</sub> criterion for four out of five test sets, as does PBE-D3 (with considerably worse statistics).

### D. Basis set and BSSE sensitivity

Assessments of DFT techniques in Secs. III B and III C demonstrate that different functionals excel in conjunction with the aug-cc-pVDZ or aug-cc-pVTZ basis set. Since certain of these results defy conventional expectations of convergence toward the Kohn–Sham limit upon expansion

of basis set size and application of the counterpoise correction, further analysis is warranted. Table VII presents MAD figures for the S22 test set computed with six hierarchical Dunning bases (cc-pVDZ, heavy-aug-cc-pVDZ, aug-cc-pVDZ, cc-pVTZ, heavy-aug-cc-pVTZ, and aug-cc-pVTZ) in both CP-corrected and non-CP-corrected modes for five of the best-performing methods: M05-2X, B970-D2, B3LYP-D3, M06-2X, and  $\omega$ B97X-D. Similar tabulations for other error schemes and additional computational approaches may be found in the supplementary material.<sup>85</sup> Two dominant patterns of response to basis set size and composition emerge for NCI-tuned functionals. The first, denoted here as “Type I,” manifests the expected  $D\zeta \rightarrow T\zeta$  trends in error statistics for hydrogen-bonded and dispersion-bound subsets, with both converging toward the low-error values achieved in the CP-corrected mode. Most functionals, including BP86-D2, B2PLYP-D3, B3LYP-D3, and  $\omega$ B97X-D belong to this group. [As do PBE0-D2 and PBE-D3, though both are somewhat scrambled since neither consistently favored one DFT/DFT-D variant for all basis sets (cf. Sec. III A)]. The second, “Type II,” is less firmly characterized but features modestly superior  $D\zeta$  results (versus  $T\zeta$ ) for HB and comparatively poor performance for CP-corrected model chemistries. The functionals XYG3, M05-2X, B970-D2, B97-D3, and M06-2X follow this pattern. A representative of each classification is illustrated in Fig. 4(a) B3LYP-D3 and (b) B970-D2, though the behavior and recommendations discussed below follow from all methods. No attempt is made to relate a technique’s classification to its mathematical form or to the circumstances of its parameterization.

In both classes, basis set performance is dramatically homogenized by application of the counterpoise correction, as is apparent by comparing the upper (non-CP-corrected) and lower (CP-corrected) portions of Figs. 4(a) and 4(b). For Type I methods, overall MAD markers for all bases are aligned about a value typically  $\leq 0.3$  kcal/mol, whereas for Type II functionals, this center can range from 0.4–0.8 kcal/mol. Within a functional, cc-pVDZ and, to a lesser extent, cc-pVTZ bases, which lack diffuse functions, exhibit larger average errors, while the remaining basis sets yield good and consistent results, with heavy-aug-cc-pVDZ striking the best balance between accuracy and computational cost. One modest exception is  $\omega$ B97X-D, for which triple- $\zeta$  basis sets should be employed with dispersion-bound systems, as haDZ and aDZ show markedly larger errors. Although deviations for HB complexes are generally the greatest in magnitude among the subsets and for DD MAD the least, this is largely due to their relative absolute interaction energies (cf. Table II); these positions typically reverse for MA%D, while the same guidelines offered here hold.

Since the counterpoise correction requires multiple computations in the expensive dimer basis, it is often preferable to overcome BSSE (which the CP approach tends to overestimate) by employing a large, flexible basis set in the non-CP-corrected mode. First examining hydrogen-bonded systems, denoted by red MAD markers in Fig. 4, the overwhelming finding is the poor performance of cc-pVXZ compared to heavy-aug-cc-pVXZ and aug-cc-pVXZ in both double- and triple- $\zeta$  regimes. This is reasonably attributed to rampant in-

termonomer basis function “borrowing” within cc-pVDZ and cc-pVTZ to express the short-range hydrogen-bonding attraction, leading to wildly overbound systems. With adequate diffuse functions, as for haXZ and aXZ, the electron density of the hydrogen bonding interaction can be properly expressed within the monomer basis, and the problem is largely resolved. Among Type I functionals, HB is notably improved at the haTZ/aTZ level by 30–60% over haDZ/aDZ, yet no uncorrected model chemistry achieves the low error of the CP-corrected mode. In contrast, Type II techniques perform more equitably for haDZ, aDZ, haTZ, and aTZ sets, with double- $\zeta$  being favored moderately, and in both regimes the non-CP-corrected MAD values are superior to the corrected. Recent work by Papajak and Truhlar<sup>101</sup> has suggested that more aggressive truncations of the diffuse basis space (beyond haXZ) are possible while preserving the accuracy of thermochemical properties.

Unlike HB systems, which demand a particular composition of basis functions, dispersion-bound complexes treated by Type I methods display a more conventional dependence on basis set size. The blue DD MAD markers in Fig. 4(a) show the improved performance of all triple- $\zeta$  bases over all double- $\zeta$  and especially the near-convergence to CP-corrected values by heavy-aug-cc-pVTZ and aug-cc-pVTZ, following from better accommodation of the long-range interaction by additional valence functions and, more successfully still, by additional diffuse functions. The -D3 correction is an almost universal improvement to each S22 subset within the B3LYP-D model chemistries considered here, and the basis set patterns described in Fig. 4(a) hold true for both DFT-D variants. Considering that all NCI systems have, in general, benefitted from diffuse basis functions, it is worthwhile to note that B3LYP-D3 results show that no further advantage is obtained from doubly augmented basis sets, d-aug-cc-pVXZ, nor from their successive truncations (analogous to haXZ) of diffuse functions from hydrogen atoms. Among all classes of interactions with Type I functionals, haXZ basis sets appear somewhat preferable to the corresponding aXZ. It should be acknowledged that another factor is at play with DFT-D methods. Particularly for DD complexes, the binding energy is heavily reliant on a (basis-independent) semiempirical correction, and techniques thus may be expected to display basis-set independence commensurate with their  $s_6$  weight [compare in Table I the consistency of aDZ and aTZ recommendations for B3LYP-D and B97-D (heavily weighted) and inconsistency for B2PLYP-D and B970-D (lightly weighted)].

Type II functionals exhibit two modes of behavior for dispersion-dominated systems. Similarly to HB, non-CP-corrected treatments of DD systems generally outperform the CP-corrected versions (M06-2X is an exception). Unlike HB, where  $D\zeta$  is consistently, if slightly, preferred over  $T\zeta$ , with DD there exist functionals for which haTZ/aTZ is clearly inferior to haDZ/aDZ (IId: XYG3, M05-2X, and B970-D2) and those for which haTZ/aTZ is superior (IIc: B97-D3 and M06-2X). Among both HB and DD, the aXZ basis is marginally preferred over its corresponding haXZ. Through the basis-set comparisons of this section, the principal strengths and weaknesses of each model chemistry for classes of NCI have been identified and are clearly consistent with



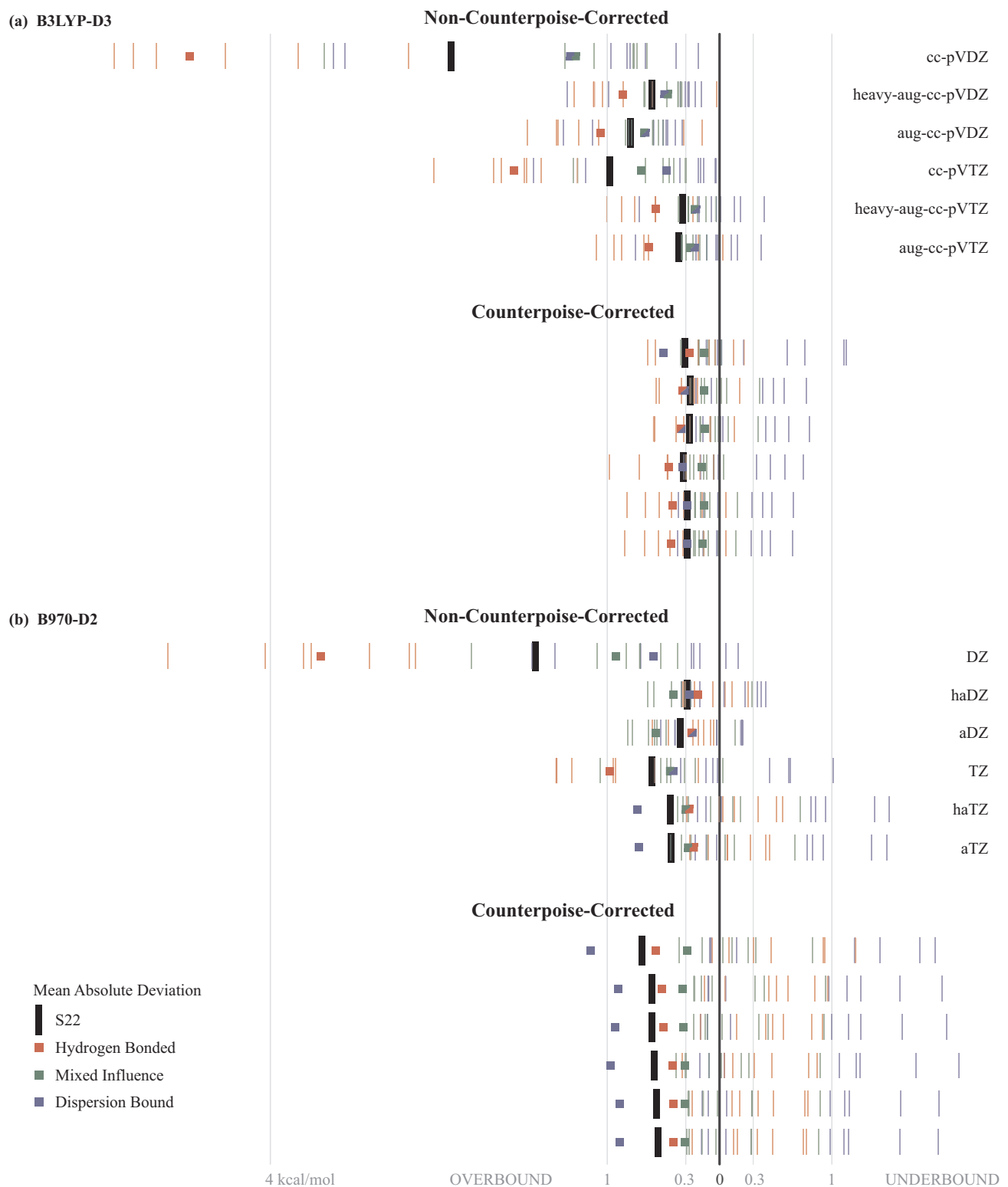


FIG. 4. Basis set and BSSE influence on S22 interaction energies. For representative DFT techniques, individual member errors and MAD summary statistics are illustrated using six hierarchical basis sets. Consult Sec. II D in the text for a concise guide to the plotting mode used here. The graph clearly demonstrates that while CP-corrected results are largely insensitive to basis set size and composition, non-CP-corrected calculations require diffuse functions for accurate treatment of hydrogen-bonded systems, and overall success for NCI is attained typically [e.g., (a) B3LYP-D3] only at triple- $\zeta$  levels of theory, though some functionals [e.g., (b) B970-D2] achieve satisfactory results with double- $\zeta$  at the cost of less-predictable patterns of error.

the recommendations offered in Table I, which were derived solely from intermethod comparisons.

The XYG3 method presents an interesting case in basis-set influence. Recently, the developers presented a study<sup>102</sup> showing a fairly steady IE error statistics upon progressively saturating the 6-311G basis with polarization and diffuse functions. In contrast, the Dunning bases chosen for this work exhibit wide variations in performance (leading to classification as Type II<sub>d</sub>). While the hydrogen-bonding subset is treated roughly consistently by four basis sets [aDZ, 6-311+G(3df,2p), 6-311++G(3df,2p), aTZ], success for dispersion-dominated systems declines precipitously for the latter two, leaving the parameterization basis, 6-311+G(3df,2p), as by far the superior of those considered. Between the two Pople basis sets, the addition of only a standard set of hydrogen-atom diffuse functions (a step not isolated in Ref. 102) induces changes of  $-0.19$  to  $+0.93$  kcal/mol in test set MAD values. Though the developers speculate on appending a “dash D”-like correction, fitting with the S22 test set indicates the optimal scaling for such a term to be essentially zero for the development basis and aDZ. That 6-311++G(3df,2p) and aTZ can achieve the same error upon scaling by  $s_6 \approx 0.25$  is further evidence of XYG3’s sensitivity with regard to dispersion and basis set composition.

## E. Examination of computational efficiency

As a brief appraisal of relative efficiency for the density functionals examined in this study, a single system of moderate size [S22-7 with 30 atoms, no symmetry, and 536

(1127) basis functions with aDZ (aTZ)] was run under conditions of minimal interference to acquire timing information for two basis sets (aDZ and aTZ), various integration grids of interest [Q-Chem: SG1 (50,194)p and (100,302); NWChem: (100,1202)], and all computational methods. Since the *ab initio* suites employed in this work have such different scaling properties, total wall time results are presented separately in Fig. 5 (a) Q-Chem and (b) NWChem. Methods are ordered in accordance with anticipated duration based upon the Perdew ladder<sup>103</sup> nomenclature and, for part (a), perform as expected, exhibiting escalations in cost after the two GGAs, the three hybrids, the two meta-GGA hybrids, and the LC hybrid techniques. This profile differs between basis sets (number of basis functions increases by  $\sim 2$  and cost for hybrids scales by  $\sim 10$ ) as the GGAs are significantly less and the LC-hybrid more expensive than the hybrids in the T $\zeta$  regime. Apart from  $\omega$ B97X-D and B3LYP-XDM, which are, respectively, insensitive and highly sensitive to grid size, functionals show the upgrade from SG1 to (100,302) to be more costly for aDZ ( $1.5$ – $2.0\times$ ) than for aTZ ( $1.1$ – $1.5\times$ ), suggesting that any calculation for which the effort of a triple- $\zeta$  basis set is expended should employ the more reliable grid. Of particular note is XDM, whose expense does not rise so quickly with the larger basis, so that a highly accurate B3LYP-XDM/aTZ job requires less than half as long again as the underlying B3LYP-D. In contrast, employing  $\omega$ B97X-D essentially doubles the B3LYP-D cost in both regimes (and doesn’t yield good results until T $\zeta$ ).

Examining functional timings results for NWChem depicted in Fig. 5(b) shows an  $\sim$ thirteenfold difference between

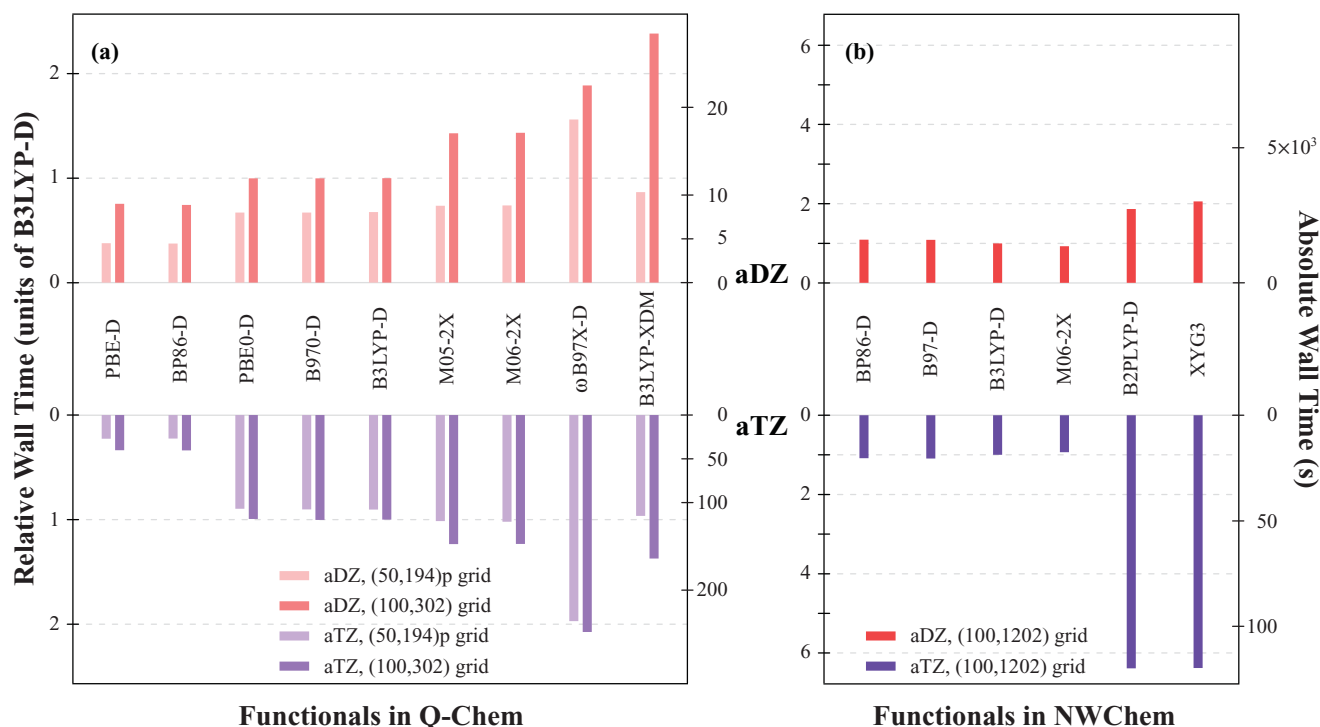


FIG. 5. Efficiency results for functionals. The total wall time relative to a B3LYP-D calculation is depicted for a number of functionals run in (a) Q-Chem and (b) NWChem, with the indicated qualities of numerical integration grid and two basis sets, aug-cc-pVDZ (upper portion) and aug-cc-pVTZ (lower portion). These results derive from a single system, S22-7, such that the basis sets differ in size by approximately a factor of 2. The right-hand axis conveys the absolute wall time. Whereas functionals in the double- $\zeta$  regime exhibit only a threefold span in relative cost, the triple- $\zeta$  results show twentyfold changes between GGAs and double-hybrids.

aug-cc-pVDZ and aug-cc-pVTZ bases for  $O(N^4)$  methods, somewhat costlier than for Q-Chem. Within both basis set profiles, the two GGA techniques surprisingly show no efficiency gain over the hybrid B3LYP-D, nor does M06-2X show any efficiency loss. This is perhaps due to peculiarities of scaling under the conditions of this timing exercise, including the substantial differences in hardware architectures between Q-Chem and NWChem calculations (IO being more burdensome for the latter). The double-hybrid methods, B2PLYP-D and XYG3, which scale as  $O(N^5)$  due to contributions from MP2, are not overly expensive with aDZ, being roughly comparable to  $\omega$ B97X-D or B3LYP-XDM. However, neither approach offers its best performance at the D $\zeta$  level, and in conjunction with aTZ, costs soar to  $\sim 6.4\times$  that of B3LYP-D (or  $\sim 85\times$  a GGA/aDZ model chemistry in Q-Chem), significantly limiting their applicability.

#### IV. SUMMARY AND CONCLUSIONS

The burgeoning development of inexpensive DFT-based approaches to nonbonding molecular complexes makes beneficial a thorough assessment of their performance and applicability. In this work, the S22 (revised in Ref. 88), HSG, and JSCH test sets of independent structures and NBC10 and HBC6 test sets of dissociation curves have been computed by a variety of functionals that have been adapted for NCI by means of DFT-D2, DFT-D3, special parameterization, or XDM. Through comparison with available CCSD(T)/CBS benchmarks, the performance of DFT methods may be assessed both overall and for hydrogen-bonded, mixed-influence, and dispersion-bound subsets. Complete evaluations of test sets have been conducted with aug-cc-pVDZ and aug-cc-pVTZ, and closer analysis of S22 with families of Dunning double- $\zeta$  and triple- $\zeta$  basis sets have permitted appraisal of these bases as well as the utility of the counterpoise correction. A tallying of relative efficiency for each of the methods employed and examination of system-size scaling has provided further guidance. Based upon the hoard of information generated in this study, the following recommendations may be made for a hydrogen-bonded, dispersion-bound, or unspecified NCI system. Note that these conclusions are suitable for molecular interactions, and interfacial/extended material contacts require additional validation. Also, functionals have undergone varying degrees of examination in the literature for other molecular properties and for general thermochemical performance.

##### *Hydrogen-bonded*

- (1) For a non-CP-corrected computation on a hydrogen-bonded system, use a basis set augmented with diffuse functions. When employing a Type I functional (see item 10), the following bases are acceptable, in order of decreasing recommendation; when employing a Type II functional, the reverse ordering is generally more appropriate:

$$\text{haTZ} \approx \text{aTZ} > \text{haDZ} \approx \text{aDZ}.$$

- (2) For a hydrogen-bonded system with a double- $\zeta$  basis set, avoid PBE-D, PBE0-D, and B2PLYP-D methods,

as the -D correction generally disrupts the fitter performance of the underlying functional. Regardless, only the last functional actually yields acceptable HB results at any basis set level.

- (3) For a hydrogen-bonded system with a double- $\zeta$  basis set, the following methods are acceptable, in order of decreasing recommendation:

$$\text{B970-D2} \approx \text{B97-D3} \approx \text{M06-2X} > \text{M05-2X} \approx \text{B2PLYP} \\ \approx \text{XYG3}.$$

- (4) For a hydrogen-bonded system, a triple- $\zeta$  basis set is not necessarily superior to a double- $\zeta$  choice. Nevertheless, the following T $\zeta$  methods are acceptable, in order of decreasing recommendation:

$$\text{B970-D2} \approx \omega\text{B97X-D} > \text{B2PLYP-D3} \approx \text{B97-D3} \\ > \text{B3LYP-D3} \approx \text{XYG3/6-311+G(3df,2p)} \\ \approx \text{M06-2X}.$$

##### *Dispersion-dominated*

- (5) For a non-CP-corrected computation on a dispersion-bound system employing a Type I functional (see item 10), use a triple- $\zeta$  basis set, where haTZ slightly outperforms aTZ. The following bases are acceptable, in order of decreasing recommendation:

$$\text{haTZ} \approx \text{aTZ} > \text{TZ}.$$

- (6) For a dispersion-bound system with a double- $\zeta$  basis set, upgrading to a triple- $\zeta$  basis for better accuracy is urged. In planning a multiple-basis study, note the lack of overlap between this list and item 7. Nevertheless, the following D $\zeta$  methods are acceptable, in order of decreasing recommendation:

$$\text{M05-2X} \approx \text{B970-D2} \approx \text{PBE0-D2} \approx \text{B97-D3} \\ > \text{B3LYP-D3} \approx \text{XYG3} \approx \text{PBE-D3}.$$

- (7) For a dispersion-bound system with a triple- $\zeta$  basis set, the following methods are acceptable, in order of decreasing recommendation:

$$\text{XYG3/6-311+G(3df,2p)} \\ \approx \text{B3LYP-D3} \approx \text{B2PLYP-D3} > \omega\text{B97X-D} \\ \approx \text{B97-D3} > \text{M06-2X} \approx \text{BP86-D2}.$$

##### *General noncovalent*

- (8) The application of Grimme's dispersion correction is decidedly helpful for noncovalently bound systems. Although both DFT-D2 and DFT-D3 variants are superior to the underlying functional (see items 2 and 9 for caveats), the following combinations are recommended:

$$\text{D2: B970, BP86, PBE0}$$

$$\text{D3: B2PLYP, B3LYP, B97, PBE}.$$

- (9) For a triple- $\zeta$  basis set, the M05-2X and M06-2X functionals can be improved in terms of error and qualitative behavior by appending the -D3 dispersion correction.
- (10) Dispersion-inclusive DFT techniques regularly follow patterns of basis-set dependence that adhere to traditional rules of convergence with respect to completeness and counterpoise correction; guidelines for these Type I methods are outlined in further items. On the other hand, functionals that elicit first-rate performance out of small basis sets often do so at the cost of sensible trends. For such Type II approaches, characterization is more fractured, and each model chemistry should be validated individually. Of the functionals investigated in this work, the following are clearly typed:

TI: BP86-D2, B2PLYP-D3, B3LYP-D3,  $\omega$ B97X-D

TII: M05-2X, B970-D2, XYG3, M06-2X, B97-D3.

- (11) For a non-CP-corrected computation on a multifunctionalized system employing a Type I functional (see item 10), use an augmented triple- $\zeta$  basis set that is safe for all noncovalent interactions. The following bases are acceptable, in order of decreasing recommendation:

haTZ  $\approx$  aTZ.

- (12) For a noncovalently interacting system with a double- $\zeta$  basis set, note that better accuracy is available with triple- $\zeta$  model chemistries (though not necessarily by applying aTZ to the following functionals, compare with item 13). Nevertheless, the following methods are acceptable, in order of decreasing recommendation:

B970-D2  $\approx$  B97-D3  $\approx$  M05-2X  $>$  M06-2X  $\approx$  XYG3.

- (13) For a noncovalently interacting system with a triple- $\zeta$  basis set, the following methods are acceptable, in order of decreasing recommendation:

B2PLYP-D3  $\approx$   $\omega$ B97X-D  $>$  B97-D3  $\approx$  B3LYP-D3  
 $\approx$  XYG3/6-311+G(3df,2p)  $>$  B970-D2  
 $\approx$  M06-2X.

- (14) When employing the Boys–Bernardi counterpoise correction, an equally accurate result may be obtained from any augmented basis set, with heavy-aug-cc-pVDZ presenting the best balance between efficiency and performance. Note that for some functionals, essentially those listed as Type II in item 10, uncorrected calculations yield lower errors. The following bases are acceptable, in order of decreasing recommendation:

haDZ  $\approx$  aDZ  $\approx$  haTZ  $\approx$  aTZ  $>$  TZ.

- (15) Balancing computational efficiency and accuracy for a noncovalently interacting system, the following combinations are worth their cost, in order of decreasing recommendation:

D $\zeta$ : B970-D2  $\approx$  B97-D3  $>$  M05-2X  $>$  M06-2X

T $\zeta$ : B97-D3  $>$  B3LYP-D3  $>$  B970-D2

$\approx$   $\omega$ B97X-D  $>$  M06-2X.

- (16) If stranded on a desert island with a hand-crank powered computer, the following single model chemistry is recommended salvaged from the luggage:

B97-D3/aug-cc-pVDZ.

Should rescue be contingent on the accuracy of a NCI calculation, the following model chemistry is recommended:

$\omega$ B97X-D/aug-cc-pVTZ.

## ACKNOWLEDGMENTS

This work was performed under the auspices of grants provided by the United States National Science Foundation (Grant No. CHE-1011360). The Center for Computational Molecular Science and Technology is funded through a NSF CRIF award (Grant No. CHE-0946869) and by Georgia Institute of Technology. B.G.S. acknowledges support from the Center for Nanophase Materials Sciences, which is sponsored at Oak Ridge National Laboratory by the Scientific User Facilities Division, U.S. Department of Energy. Á.V.-M. acknowledges support from the Department of Energy, Offices of Basic Energy Science and Advanced Scientific Computing Research as part of the SciDAC program. This research used resources supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725 and advanced computing resources provided by the National Science Foundation. The computations were performed partially on Kraken at the National Institute for Computational Sciences (<http://www.nics.tennessee.edu/>). The authors wish to thank Professor Stefan Grimme for helpful discussions on the DFT-D3 method and Mr. Edward G. Hohenstein for inspiration regarding a figure.

<sup>1</sup>E. A. Meyer, R. K. Castellano, and F. Diederich, *Angew. Chem., Int. Ed. Engl.* **42**, 1210 (2003).

<sup>2</sup>A. Anbarasu, S. Anand, M. M. Babu, and R. Sethumadhavan, *Int. J. Biol. Macromol.* **41**, 251 (2007).

<sup>3</sup>G. A. Jeffrey and W. Saenger, *Hydrogen Bonding in Biological Structures* (Springer-Verlag, Berlin, 1991).

<sup>4</sup>L. Brunsfeld, B. J. B. Folmer, E. W. Meijer, and R. P. Sijbesma, *Chem. Rev.* **101**, 4071 (2001).

<sup>5</sup>C. G. Claessens and J. F. Stoddart, *J. Phys. Org. Chem.* **10**, 254 (1997).

<sup>6</sup>T. C. Dinadayalane, L. Gorb, T. Simeon, and H. Dodziuk, *Int. J. Quantum Chem.* **107**, 2204 (2007).

<sup>7</sup>T. Takatani, E. G. Hohenstein, and C. D. Sherrill, *J. Chem. Phys.* **128**, 124111 (2008).

<sup>8</sup>S. Grimme, *J. Chem. Phys.* **118**, 9095 (2003).

<sup>9</sup>R. A. Distasio and M. Head-Gordon, *Mol. Phys.* **105**, 1073 (2007).

<sup>10</sup>A. Tkatchenko, R. A. DiStasio, Jr., M. Head-Gordon, and M. Scheffler, *J. Chem. Phys.* **131**, 094106 (2009).

<sup>11</sup>M. Pitoňák, P. Neogrády, J. Černý, S. Grimme, and P. Hobza, *ChemPhysChem* **10**, 282 (2009).

<sup>12</sup>B. Jeziorski, R. Moszynski, and K. Szalewicz, *Chem. Rev.* **94**, 1887 (1994).

<sup>13</sup>A. Heßelmann, G. Jansen, and M. Schütz, *J. Chem. Phys.* **122**, 014103 (2005).

<sup>14</sup>E. G. Hohenstein and C. D. Sherrill, *J. Chem. Phys.* **133**, 104107 (2010).

<sup>15</sup>O. A. von Lilienfeld, I. Tavernelli, U. Rothlisberger, and D. Sebastiani, *Phys. Rev. Lett.* **93**, 153004 (2004).



- <sup>16</sup>Q. Wu and W. Yang, *J. Chem. Phys.* **116**, 515 (2002).
- <sup>17</sup>S. Grimme, *J. Comput. Chem.* **25**, 1463 (2004).
- <sup>18</sup>S. Grimme, *J. Comput. Chem.* **27**, 1787 (2006).
- <sup>19</sup>S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, *J. Chem. Phys.* **132**, 154104 (2010).
- <sup>20</sup>A. D. Becke and E. R. Johnson, *J. Chem. Phys.* **122**, 154104 (2005).
- <sup>21</sup>A. D. Becke and E. R. Johnson, *J. Chem. Phys.* **124**, 014104 (2006).
- <sup>22</sup>J. G. Angyan, *J. Chem. Phys.* **127**, 024108 (2007).
- <sup>23</sup>A. D. Becke and E. R. Johnson, *J. Chem. Phys.* **127**, 154108 (2007).
- <sup>24</sup>A. Heßelmann, *J. Chem. Phys.* **130**, 084104 (2009).
- <sup>25</sup>S. N. Steinmann and C. Corminboeuf, *J. Chem. Theory Comput.* **6**, 1990 (2010).
- <sup>26</sup>X. Xu and W. A. Goddard, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 2673 (2004).
- <sup>27</sup>Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A* **109**, 5656 (2005).
- <sup>28</sup>Y. Zhao, N. E. Schultz, and D. G. Truhlar, *J. Chem. Theory Comput.* **2**, 364 (2006).
- <sup>29</sup>Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.* **120**, 215 (2008).
- <sup>30</sup>J. Chai and M. Head-Gordon, *Phys. Chem. Chem. Phys.* **10**, 6615 (2008).
- <sup>31</sup>S. Grimme, *J. Chem. Phys.* **124**, 034108 (2006).
- <sup>32</sup>T. Schwabe and S. Grimme, *Phys. Chem. Chem. Phys.* **9**, 3397 (2007).
- <sup>33</sup>J. Chai and M. Head-Gordon, *J. Chem. Phys.* **131**, 174105 (2009).
- <sup>34</sup>Y. Zhang, X. Xu, and W. A. Goddard, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 4963 (2009).
- <sup>35</sup>D. C. Langreth, M. Dion, H. Rydberg, E. Schroder, P. Hyldgaard, and B. I. Lundqvist, *Int. J. Quantum Chem.* **101**, 599 (2005).
- <sup>36</sup>O. A. Vydrov and T. Van Voorhis, *Phys. Rev. Lett.* **103**, 063004 (2009).
- <sup>37</sup>C. D. Sherrill, T. Takatani, and E. G. Hohenstein, *J. Phys. Chem. A* **113**, 10146 (2009).
- <sup>38</sup>M. Pitoňák, K. E. Riley, Neogrady, and P. Hobza, *ChemPhysChem* **9**, 1636 (2008).
- <sup>39</sup>E. C. Lee, D. Kim, P. Jurečka, P. Tarakeshwar, P. Hobza, and K. S. Kim, *J. Phys. Chem. A* **111**, 3446 (2007).
- <sup>40</sup>K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, *Chem. Phys. Lett.* **157**, 479 (1989).
- <sup>41</sup>T. J. Lee and G. E. Scuseria, "Achieving chemical accuracy with coupled-cluster theory," in *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*, edited by S. R. Langhoff (Kluwer Academic Publishers, Dordrecht, 1995), pp. 47–108.
- <sup>42</sup>A. K. Rappé and E. R. Bernstein, *J. Phys. Chem. A* **104**, 6117 (2000).
- <sup>43</sup>S. Tsuzuki and H. P. Lüthi, *J. Chem. Phys.* **114**, 3949 (2001).
- <sup>44</sup>E. R. Johnson, R. A. Wolkow, and G. A. DiLabio, *Chem. Phys. Lett.* **394**, 334 (2004).
- <sup>45</sup>F. O. Kannemann and A. D. Becke, *J. Chem. Theory Comput.* **5**, 719 (2009).
- <sup>46</sup>W. Koch and M. C. Holthausen, *A Chemist's Guide to Density Functional Theory* (Wiley-VCH, New York, 2001).
- <sup>47</sup>Y. Zhao and D. G. Truhlar, *Acc. Chem. Res.* **41**, 157 (2008).
- <sup>48</sup>E. R. Johnson, I. D. Mackie, and G. A. DiLabio, *J. Phys. Org. Chem.* **22**, 1127 (2009).
- <sup>49</sup>J. Gräfenstein and D. Cremer, *J. Chem. Phys.* **130**, 124105 (2009).
- <sup>50</sup>C. D. Sherrill, "Computations of noncovalent  $\pi$  interactions," in *Reviews in Computational Chemistry*, edited by K. B. Lipkowitz and T. R. Cundari (Wiley, Hoboken, 2009), Vol. 26, pp. 1–38.
- <sup>51</sup>M. E. Foster and K. Sohlberg, *Phys. Chem. Chem. Phys.* **12**, 307 (2010).
- <sup>52</sup>K. E. Riley, M. Pitoňák, P. Jurečka, and P. Hobza, *Chem. Rev.* **110**, 5023 (2010).
- <sup>53</sup>Y. Zhao, B. J. Lynch, and D. G. Truhlar, *J. Phys. Chem. A* **108**, 4786 (2004).
- <sup>54</sup>J. Antony and S. Grimme, *Phys. Chem. Chem. Phys.* **8**, 5287 (2006).
- <sup>55</sup>P. Jurečka, J. Šponer, J. Černý, and P. Hobza, *Phys. Chem. Chem. Phys.* **8**, 1985 (2006).
- <sup>56</sup>P. Jurečka, J. Cerny, P. Hobza, and D. R. Salahub, *J. Comput. Chem.* **28**, 555 (2007).
- <sup>57</sup>E. G. Hohenstein, S. T. Chill, and C. D. Sherrill, *J. Chem. Theory Comput.* **4**, 1996 (2008).
- <sup>58</sup>K. E. Riley, M. Pitoňák, J. Cerny, and P. Hobza, *J. Chem. Theory Comput.* **6**, 66 (2010).
- <sup>59</sup>A. Vazquez-Mayagoitia, C. D. Sherrill, E. Apra, and B. G. Sumpter, *J. Chem. Theory Comput.* **6**, 727 (2010).
- <sup>60</sup>K. S. Thanthiriwat, E. G. Hohenstein, L. A. Burns, and C. D. Sherrill, *J. Chem. Theory Comput.* **7**, 88 (2011).
- <sup>61</sup>L. Gráfová, M. Pitoňák, J. Řezáč, and P. Hobza, *J. Chem. Theory Comput.* **6**, 2365 (2010).
- <sup>62</sup>F. O. Kannemann and A. D. Becke, *J. Chem. Theory Comput.* **6**, 1081 (2010).
- <sup>63</sup>M. O. Sinnokrot and C. D. Sherrill, *J. Phys. Chem. A* **108**, 10200 (2004).
- <sup>64</sup>L. F. Molnar, X. He, B. Wang, and K. M. Merz, *J. Chem. Phys.* **131**, 065102 (2009).
- <sup>65</sup>The DFT-D2  $s_6$  parameter for the B970 functional was obtained by minimizing the mean absolute percent deviation for the S22 test set calculated with uncounterpoise-corrected and equally weighted aug-cc-pVDZ and aug-cc-pVTZ basis sets. A more thorough procedure, taking into account BSSE-corrected interaction energies, shifts the recommended parameter only slightly, to 0.80.
- <sup>66</sup>See <http://www.uni-muenster.de/Chemie.oc/grimme/> for a FORTRAN program implementing the DFT-D3 method and a file with available  $C_6$  coefficients. Westfälische Wilhelms-Universität Münster, Organisch-Chemisches Institut, Theoretische Organische Chemie, Grimme Research Group, Corrensstrasse 40, D-48149 Münster, email: grimmes@uni-muenster.de.
- <sup>67</sup>J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- <sup>68</sup>C. Adamo and V. Barone, *J. Chem. Phys.* **110**, 6158 (1999).
- <sup>69</sup>M. Ernzerhof and G. E. Scuseria, *J. Chem. Phys.* **110**, 5029 (1999).
- <sup>70</sup>A. D. Becke, *Phys. Rev. A* **38**, 3098 (1988).
- <sup>71</sup>J. P. Perdew, *Phys. Rev. B* **33**, 8822 (1986).
- <sup>72</sup>A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).
- <sup>73</sup>P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, *J. Phys. Chem.* **98**, 11623 (1994).
- <sup>74</sup>C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).
- <sup>75</sup>It is acknowledged that the "exact HF exchange" energy computed within a DFT method is always slightly higher than the quantity produced by HF theory, due to imprecision in representing the solution to the HF ansatz by KS orbitals. See Ch. 5.3 of Ref. 46 for further discussion. Nevertheless, the present work continues the widespread conventional use of the term.
- <sup>76</sup>A. D. Becke, *J. Chem. Phys.* **107**, 8554 (1997).
- <sup>77</sup>L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, *J. Chem. Phys.* **106**, 1063 (1997).
- <sup>78</sup>L. A. Curtiss, P. C. Redfern, K. Raghavachari, and J. A. Pople, *J. Chem. Phys.* **109**, 42 (1998).
- <sup>79</sup>Y. Zhao and D. G. Truhlar, *J. Chem. Theory Comput.* **3**, 289 (2007).
- <sup>80</sup>L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, *J. Chem. Phys.* **112**, 7374 (2000).
- <sup>81</sup>A. D. Becke and M. R. Roussel, *Phys. Rev. A* **39**, 3761 (1989).
- <sup>82</sup>J. Kong, Z. T. Gan, E. Proynov, M. Freindorf, and T. R. Furlani, *Phys. Rev. A* **79**, 042510 (2009).
- <sup>83</sup>L. A. Burns, K. S. Thanthiriwat, J. Kong, and C. D. Sherrill, "Improved performance for the exchange-hole dipole moment method: A reparameterization and assessment," (unpublished).
- <sup>84</sup>A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, and A. K. Wilson, *Chem. Phys. Lett.* **286**, 243 (1998).
- <sup>85</sup>See supplementary material at <http://dx.doi.org/10.1063/1.3545971> for Tables 1–5 elaborating the reference CBS extrapolation level for individual test set members; Table 6 defining summary error statistics; Tables 7–14 containing MA%D and MA%BD counterparts to Tables VII and III–VI; Tables 15–16 compiling MAD results for alternate DFT-D3 parameters; and a shocking number of tables detailing interaction energies. Text files with electronic energies and interaction energies are also available, along with Cartesian coordinates for all test set members.
- <sup>86</sup>Y. Zhao and D. G. Truhlar, *J. Chem. Theory Comput.* **1**, 415 (2005).
- <sup>87</sup>J. Řezáč, P. Jurečka, K. E. Riley, J. Černý, H. Valdes, K. Pluháčková, K. Berka, T. Řezáč, M. Pitoňák, J. Vondrášek, and P. Hobza, *Collect. Czech. Chem. Commun.* **73**, 1261 (2008). <http://www.begdb.com>.
- <sup>88</sup>T. Takatani, E. G. Hohenstein, M. Malagoli, M. S. Marshall, and C. D. Sherrill, *J. Chem. Phys.* **132**, 144104 (2010).
- <sup>89</sup>T. Takatani and C. D. Sherrill, *Phys. Chem. Chem. Phys.* **9**, 6106 (2007).
- <sup>90</sup>E. G. Hohenstein and C. D. Sherrill, *J. Phys. Chem. A* **113**, 878 (2009).
- <sup>91</sup>J. Faver, M. L. Benson, X. He, B. P. Roberts, B. Wang, M. S. Marshall, M. R. Kennedy, C. D. Sherrill, and K. M. Merz Jr., "Formal Estimation of Errors in Computed Absolute Interaction Energies for Protein-Ligand Complexes," *J. Chem. Theory Comput.* (in press) (2011).
- <sup>92</sup>As was verified through inspection of base pair classification and comparison of recomputed MP2/aug-cc-pVDZ interaction energies with those

- in the published table, the interaction energy and label of JSCH-40 is properly associated with the geometry in file 100, represented as JSCH-40 $\leftarrow$ 100. Similarly, JSCH-41 $\leftarrow$ 102, JSCH-42 $\leftarrow$ 41, JSCH-45 $\leftarrow$  46, JSCH-46 $\leftarrow$ 45, JSCH-99 $\leftarrow$ 40, JSCH-100 $\leftarrow$  99, and JSCH-102 $\leftarrow$ 42, for a total of eight reassignments.
- <sup>93</sup>T. H. Dunning, *J. Chem. Phys.* **90**, 1007 (1989).
- <sup>94</sup>R. A. Kendall, T. H. Dunning, and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).
- <sup>95</sup>S. F. Boys and F. Bernardi, *Mol. Phys.* **19**, 553 (1970).
- <sup>96</sup>E. J. Bylaska, W. A. de Jong, N. Govind, K. Kowalski, T. P. Straatsma, M. Valiev, D. Wang, E. Apra, T. L. Windus, J. Hammond, P. Nichols, S. Hirata, M. T. Hackler, Y. Zhao, P.-D. Fan, R. J. Harrison, M. Dupuis, D. M. A. Smith, J. Nieplocha, V. Tipparaju, M. Krishnan, Q. Wu, T. Van Voorhis, A. A. Auer, M. Nooijen, E. Brown, G. Cisneros, G. I. Fann, H. Fruchtl, J. Garza, K. Hirao, R. Kendall, J. A. Nichols, K. Tsemekhman, K. Wolinski, J. Anchell, D. Bernholdt, P. Borowski, T. Clark, D. Clerc, H. Dachsel, M. Deegan, K. Dyall, D. Elwood, E. Glendening, M. Gutowski, A. Hess, J. Jaffe, B. Johnson, J. Ju, R. Kobayashi, R. Kutteh, Z. Lin, R. Littlefield, X. Long, B. Meng, T. Nakajima, S. Niu, L. Pollack, M. Rosing, G. Sandrone, M. Stave, H. Taylor, G. Thomas, J. van Lenthe, A. Wong, and Z. Zhang, *NWChem, A Computational Chemistry Package for Parallel Computers, version 5.1* (Pacific Northwest National Laboratory, Richland, 2007).
- <sup>97</sup>Y. Shao, L. F. Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. T. Brown, A. T. B. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, R. A. DiStasio Jr., R. C. Lochan, T. Wang, G. J. O. Beran, N. A. Besley, J. M. Herbert, C. Y. Lin, T. Van Voorhis, S. H. Chien, A. Sodt, R. P. Steele, V. A. Rassolov, P. E. Maslen, P. P. Korambath, R. D. Adamson, B. Austin, J. Baker, E. F. C. Byrd, H. Dachsel, R. J. Doerksen, A. Dreuw, B. D. Dunietz, A. D. Dutoi, T. R. Furlani, S. R. Gwaltney, A. Heyden, S. Hirata, C.-P. Hsu, G. Kedziora, R. Z. Khalliulin, P. Klunzinger, A. M. Lee, M. S. Lee, W. Liang, I. Lotan, N. Nair, B. Peters, E. I. Proynov, P. A. Pieniazek, Y. M. Rhee, J. Ritchie, E. Rosta, C. D. Sherrill, A. C. Simmonett, J. E. Subotnik, H. L. Woodcock, W. Zhang, A. T. Bell, A. K. Chakraborty, D. M. Chipman, F. J. Keil, A. Warshel, W. J. Hehre, H. F. Schaefer, J. Kong, A. I. Krylov, P. M. W. Gill, and M. Head-Gordon, *Phys. Chem. Chem. Phys.* **8**, 3172 (2006).
- <sup>98</sup>E. R. Johnson, A. D. Becke, C. D. Sherrill, and G. A. DiLabio, *J. Chem. Phys.* **131**, 034111 (2009).
- <sup>99</sup>J. A. Pople, in *Energy, Structure and Reactivity: Proceedings of the 1972 Boulder Summer Research Conference on Theoretical Chemistry*, edited by D. W. Smith and W. B. McRae, (Wiley, New York, 1973) p. 51.
- <sup>100</sup>A. Karton, D. Gruzman, and J. M. L. Martin, *J. Phys. Chem. A* **113**, 8434 (2009).
- <sup>101</sup>E. Papajak and D. G. Truhlar, *J. Chem. Theory Comput.* **6**, 597 (2010).
- <sup>102</sup>I. Y. Zhang, Y. Luo, and X. Xu, *J. Chem. Phys.* **133**, 104105 (2010).
- <sup>103</sup>J. P. Perdew and K. Schmidt, *AIP Conf. Proc.* **577**, 1 (2001).